ARKIV FÖR FYSIK Band 17 nr 16

Communicated 10 February 1960 by H. Cramér and O. Klein

On the foundations of the linear theory of irreversible processes. II

By U. Uhlhorn

1. Introduction

In a previous paper [Uhlhorn, 1960 I][1] some aspects of the phenomenological linear theory of irreversible processes were discussed. The present work is concerned with the probabilistic interpretation of the phenomenological theory and its relation to statistical mechanics.

In the first sections of the paper a mathematical framework suitable for the statistical description of irreversible processes is presented. This framework is a somewhat extended version of the Hilbert space theory of stationary stochastic processes. Its formal structure is similar to that of quantum mechanics, and this analogy turns out to be fruitful (sections 5, 8, Appendix 2). The classification of a stationary stochastic process as deterministic or non deterministic—these concepts were originally introduced by Wold [Wold, 1938]—corresponds closely to the classification of a physical process as reversible or irreversible (sections 8, 10, 13). A non deterministic stochastic process, which corresponds to an irreversible physical process, can always be considered as generated by a "purely random" mechanism. The connection between the stochastic process and its generating mechanism is expressed in the form of a "causal representation" of the stochastic process (sections 9, 10, 13). In sections 11 and 12 the application of the linear theory of stochastic processes to statistical mechanics is discussed, and section 13 is devoted to the probabilistic interpretation of the linear phenomenological theory of irreversible processes.

2. Some basic probabilistic concepts

In this section a probabilistic framework adapted for the treatment of irreversible processes is presented. We shall be concerned with the so called linear theory of stochastic processes[2] in a slightly modified and extended form.

The basic concept of the theory of probability is that of an "event" and its assigned "probability". Conventionally the events are represented as certain subsets of a basic "phase space" Ω. It is assumed that the complement of an event is an event, that the empty set ϕ is an event, and that the logical sum of enumerably many events is an event. The totality of all events $\mathfrak{F}$ is then said to be a sigma-field. In general a point in Ω is not an event. To every event F there is assigned a probability $P(F)$ that is a non negative number in such a way that $P(\Omega)=1$ and that the probability of the logical sum of denumerably many incompatible events is equal to the sum of the probabilities of the events. $P(.)$ is then said to be a sigma-additive, normalized measure on $(\Omega, \mathfrak{F})$, and the structure $(\Omega, \mathfrak{F}, P(.))$ is said to be a probability space.

---

[1] Hereafter referred to as I.

[2] See in particular [Karhunen, 1947] and [Grenander and Szegö, 1958]. References to later publications on the subject can be found in [Jaglom, 1959].

A random variable $x=x(.)$ is a real or complex function on Ω, such that if I is an interval of the real axis or of the complex plane the subset $x^{-1}I$ of Ω, consisting of all points ω for which the value $x(\omega)$ is contained in I, is an event. Hence the probability distribution of x on the real axis or on the complex plane is well defined by $P_x(I)=P(x^{-1}I)$.

The expectation or mean value of a random variable x is defined as the integral

$$\mathbf{E}x=\int_\Omega x(\omega)P(d\omega)=\int x' P_x(dx') \text{ if } \mathbf{E}|x|<\infty.$$

If $\mathbf{E}|x|^2<\infty$ the random variable x is said to have finite variance. The totality of all complex random variables $x, y, \ldots$ with finite variance form a complex Hilbert space $L^2(\Omega, \mathfrak{F}, P(.))$ with the scalar product

$$<x|y> = \int_\Omega \overline{x(\omega)}\, y(\omega)\, P(d\omega)^1 \tag{1}$$

and the norm

$$\|x\|=[\int_\Omega |x(\omega)|^2 P(d\omega)]^{1/2}. \tag{2}$$

If $\mathbf{E}x=\mathbf{E}y=0$ the scalar product (1) is called the covariance of the random variables x and y. We note that $e(\omega)\equiv 1$ defines a random variable e with finite variance, $\|e\|=1$. The expectation of an arbitrary random variable x with finite variance can thus be written

$$\mathbf{E}x=e<e|x>, \tag{3}$$

and it is seen that $\mathbf{E}$ may be considered as a projection on $L^2(\Omega, \mathfrak{F}, P(.))$ onto the subspace of constant random variables spanned by the unit vector e.

For the purposes of statistical mechanics of closed classical systems it is always possible to take the basic space Ω as the phase space of the canonical coordinates. All intervals of the phase space are considered as events. In the statistical mechanics of macroscopic systems every physically relevant probability distribution is absolutely continuous with respect to the "Liouville measure" const. $dp_1 \ldots dp_n\, dq_1 \ldots dq_n$, that is it is characterized by a density $f(p_1, \ldots, p_n, q_1, \ldots, q_n)$. If the probability assigned to an event is zero the event is considered as unobservable. Events consisting of a single phase point are thus always unobservable. For quantum mechanical systems a correspond-

---

[1] The bar denotes complex conjugation.

ing phase space can be constructed relative to a particular series of observations [Uhlhorn, 1960 b].

The statistical description of open systems can in general not be given in terms of a probability distribution on the phase space only. If for instance we consider a system the dynamics of which is determined by Hamilton's equations combined with a random force acting at the boundary we will in general need an infinite dimensional basic space Ω to give a complete probabilistic characterization of the system and the changes taking place in it in the course of time even though the number of canonical coordinates describing the system at an arbitrary instant is finite.

3. Stochastic processes with finite variance

A stochastic process is a family of random variables. We shall consider families of random variables with finite variance exclusively. It is convenient to characterize a stochastic process as a random function: Let $\mathcal{T}$ be an arbitrary set and associate with each point τ in $\mathcal{T}$ a random variable $x(\tau)$. $\mathcal{T}$ is called the parameter set of the stochastic process $x(.)$. For every fixed point ω in $\Omega, x(., \omega)$ is a function of τ, a realization of the stochastic process.

The function $r(\tau_1, \tau_2)$ defined by

$$r(\tau_1, \tau_2) = \overline{r(\tau_2, \tau_1)} = \langle x(\tau_1) \,|\, x(\tau_2) \rangle \tag{4}$$

is called the covariance function of the stochastic process $\{x(\tau) \,|\, \tau \in \mathcal{T}\}$.

Let $G_0(\mathcal{T})$ be the set of all complex functions $\varphi(.), \psi(.), \ldots$ on the parameter space $\mathcal{T}$ which are zero in all but a finite number of points. Evidently $G_0(\mathcal{T})$ is a complex vector space. Let a linear transformation from $G_0(\mathcal{T})$ into $L^2(\Omega, \mathfrak{F}, P(.))$ be defined by

$$x[\varphi] = \sum_{\tau \in \mathcal{T}} x(\tau)\, \varphi(\tau). \tag{5}$$

We have obtained a new stochastic process $x[.]$ with the parameter set $G_0(\mathcal{T})$ which in contrast to the original process $x(.)$ is a linear function of its argument. Evidently the two processes $x(.)$ and $x[.]$ determine each other uniquely. The covariance function corresponding to (5) is

$$\mathcal{R}(\varphi, \psi) = \langle x[\varphi] \,|\, x[\psi] \rangle = \sum_{\tau_1, \tau_2 \in \mathcal{T}} \sum r(\tau_1, \tau_2)\, \overline{\varphi(\tau_1)}\, \psi(\tau_2), \tag{6}$$

and is a Hermitian non negative functional on the vector space $G_0(\mathcal{T})$

$$\begin{aligned} &\mathcal{R}(\varphi, \psi_1 + \psi_2) = \mathcal{R}(\varphi, \psi_1) + \mathcal{R}(\varphi, \psi_2), \ \mathcal{R}(\varphi, c\psi) = c\,\mathcal{R}(\varphi, \psi) \\ &\overline{\mathcal{R}(\varphi, \psi)} = \mathcal{R}(\psi, \varphi) \\ &\mathcal{R}(\varphi, \varphi) \geqslant 0.^{1} \end{aligned} \tag{7}$$

---

[1] The third property evidently implies the second one as the vector space $G_0(\mathcal{T})$ is complex.

The range of $x[.]$, that is the totality of random variables $x[\varphi]$ with φ in $G_0(\mathcal{J})$ is a linear space contained in $L^2(\Omega, \mathcal{F}, P(.))$. The closure[1] of this range $H(\mathcal{J})$ is called the Hilbert space of the stochastic process $x(.)$.

Starting from any functional $\mathcal{R}(\varphi, \psi)$ with the properties (7) one can easily construct a Hilbert space $G(\mathcal{J})$ which is isometrically equivalent to the space $H(\mathcal{J})$, but is entirely "non probabilistic". To see this we define for fixed ψ an antilinear functional f_ψ on $G_0(\mathcal{J})$ by

$$f_\psi[\varphi] = \mathcal{R}(\varphi, \psi). \tag{8}$$

The totality of functionals of the type (8) form a linear complex space

$$f_{\psi+\varphi} = f_\psi + f_\varphi, \quad f_{c\psi} = c f_\psi, \tag{9}$$

and

$$f_\psi = 0 \text{ if and only if } \mathcal{R}(\psi, \psi) = 0. \tag{10}$$

The last proposition follows from the fact that

$$|\mathcal{R}(\varphi, \psi)|^2 \leqslant \mathcal{R}(\varphi, \varphi)\, \mathcal{R}(\psi, \psi). \tag{11}$$

Hence a scalar product can be uniquely defined by

$$\langle f_\varphi | f_\psi \rangle = \mathcal{R}(\varphi, \psi), \quad \| f_\varphi \| = \mathcal{R}(\varphi, \varphi)^{\frac{1}{2}}, \tag{12}$$

and closing the space of antilinear functionals on $G_0(\mathcal{J})$ with respect to the norm (12) we obtain a Hilbert space $G(\mathcal{J})$. Hence it is always possible to construct a Hilbert space corresponding to a prescribed non negative functional of the type (7). Evidently there is a one to one correspondence between the random variable $x[\varphi]$ and the functional f_φ and $\|x[\varphi]\| = \|f_\varphi\|$. Consequently the Hilbert spaces $H(\mathcal{J})$ and $G(\mathcal{J})$ are isometrically equivalent.

4. Linear conditioning

Let us assume that $\mathcal{J}_1$ is a subset of $\mathcal{J}$ and that the process $x(.)$ has been observed for all values of τ in $\mathcal{J}_1$. This means that we have obtained a particular realization $\{x(\tau, \omega_0) | \tau \in \mathcal{J}_1\}$ of the stochastic process. This knowledge given, we want to estimate the value of the process for an arbitrary value of the parameter τ. Thus we need some sort of conditional expectation of $x(\tau)$ given the realization $\{x(\tau_1) | \tau_1 \in \mathcal{J}_1\}$. Such conditional expectations can be constructed in various ways. We shall use linear conditioning, that is we shall choose an estimate $x(\tau, \mathcal{J}_1)$ of $x(\tau)$ which is a linear combination of random variables $x(\tau_1)$ with parameter values in $\mathcal{J}_1$ or a limit of such combinations. This means that we take $x(\tau, \mathcal{J}_1)$ to be an element in $H(\mathcal{J}_1)$. The estimation error is $x(\tau) - x(\tau, \mathcal{J}_1)$, and we want to determine $x(\tau, \mathcal{J}_1)$ so that the variance of the estimation error $\|x(\tau) - x(\tau, \mathcal{J}_1)\|^2$ is as small as possible. It is geometrically evident that $x(\tau, \mathcal{J}_1)$ has to be taken as the orthogonal projection of $x(\tau)$ into $H(\mathcal{J}_1)$

$$x(\tau, \mathcal{J}_1) = E^{\mathcal{J}_1} x(\tau) \quad (E^{\mathcal{J}_1} \text{ is the projection onto } H(\mathcal{J}_1)). \tag{13}$$

---

[1] The closure is taken with respect to the norm (2); for the definition see e.g. [Halmos, 1951].

It can evidently happen that

$$E^{\mathcal{J}_1} x(\tau) = x(\tau) \tag{14}$$

even if τ is not contained in $\mathcal{J}_1$. Then $x(\tau)$ is said to be *determined* by the process in $\mathcal{J}_1$.

It is obvious that the solution (13) of the estimation problem is purely formal. To obtain the real solution one needs an analytical characterization of the Hilbert subspace $H(\mathcal{J}_1)$.

5. Operations on stochastic processes

Let a be a point transformation on the parameter space $\mathcal{J}$ of the stochastic process $x(.)$, that is for every τ in $\mathcal{J}$, $a\tau$ is a point in $\mathcal{J}$. We define the linear transformation $\mathcal{D}[a]$ corresponding to a in the linear space $G_0(\mathcal{J})$ by

$$\mathcal{D}[a]\varphi(\tau) = \sum_{a\tau' = \tau} \varphi(\tau'). \tag{15}$$

It follows directly from the definition (15) that if a_1 and a_2 are two point transformations in $\mathcal{J}$ we have

$$\mathcal{D}[a_2]\mathcal{D}[a_1] = \mathcal{D}[a_2 a_1]. \tag{16}$$

We shall further assume that the transformation a is bounded in the following sense:

$$\mathcal{R}(\mathcal{D}[a]\varphi, \mathcal{D}[a]\varphi) = \sum_{\tau_1}\sum_{\tau_2} r(a\tau_1, a\tau_2)\overline{\varphi(\tau_1)}\varphi(\tau_2) \leqslant M_a^2 \mathcal{R}(\varphi, \varphi) \tag{17}$$

for every element φ in $G_0(\mathcal{J})$ and some non negative constant M_a. It is a direct consequence of (17) that

$$\mathcal{R}(\varphi, \varphi) = 0 \text{ implies } \mathcal{R}(\mathcal{D}[a]\varphi, \mathcal{D}[a]\varphi) = 0. \tag{18}$$

Having obtained (18) we may without ambiguity define a linear operation $D[a]$ corresponding to a in $H(\mathcal{J})$. We begin by defining $D[a]$ for elements of the type $x[\varphi]$:

$$D[a]x[\varphi] = x[\mathcal{D}[a]\varphi]. \tag{19}$$

The definition (19) is unique for if $x[\varphi] = x[\psi]$ we have according to (18)

$$x[\mathcal{D}[a]\varphi] = x[\mathcal{D}[a]\psi].$$

It follows from (17) that

$$\| D[a]x[\varphi] \| \leqslant M_a \| x[\varphi] \|. \tag{20}$$

Hence $D[a]$ may be extended to the complete Hilbert space $H(\mathcal{J})$, and

$$\| D[a] \| \leqslant M_a. \tag{21}$$

It follows from (19) by taking $\varphi(\tau') = \begin{cases} 1\,(\tau' = \tau) \\ 0\,(\tau' \neq \tau) \end{cases}$ that

$$D[a]\,x(\tau) = x(a\,\tau). \tag{22}$$

If a is a point transformation in $\mathfrak{J}$ satisfying the condition (17) and if a^* is a point transformation in $\mathfrak{J}$ related to a by the condition on the covariance function

$$r(\tau_1, a\,\tau_2) = r(a^*\tau_1, \tau_2) \quad (\text{arbitrary } \tau_1, \tau_2 \in \mathfrak{J}), \tag{23}$$

then $D[a^*]$ exists and is equal to the Hermitian adjoint of $D[a]$:

$$D[a^*] = D[a]^*. \tag{24}$$

To see this we observe that (23) implies

$$\mathcal{R}(\varphi, \mathcal{D}[a]\,\psi) = \mathcal{R}(\mathcal{D}[a^*]\,\varphi, \psi) \quad (\text{arbitrary } \varphi, \psi \in G_0(\mathfrak{J})). \tag{25}$$

Applying Schwarz' inequality we find

$$\begin{aligned} \mathcal{R}(\mathcal{D}[a^*]\,\varphi, \mathcal{D}[a^*]\,\varphi) &= \mathcal{R}(\varphi, \mathcal{D}[a]\mathcal{D}[a^*]\,\varphi) \leqslant \\ &\leqslant \mathcal{R}(\varphi, \varphi)^{1/2}\,\mathcal{R}(\mathcal{D}[a]\mathcal{D}[a^*]\,\varphi, \mathcal{D}[a]\mathcal{D}[a^*]\,\varphi)^{1/2} \leqslant \\ &\leqslant \mathcal{R}(\varphi, \varphi)^{1/2}\,M_a\,\mathcal{R}(\mathcal{D}[a^*]\,\varphi, \mathcal{D}[a^*]\,\varphi)^{1/2}, \end{aligned} \tag{26}$$

and hence

$$\mathcal{R}(\mathcal{D}[a^*]\,\varphi, \mathcal{D}[a^*]\,\varphi) \leqslant M_a^2\,\mathcal{R}(\varphi, \varphi). \tag{27}$$

It follows that $D[a^*]$ is well defined and that

$$\langle D[a^*]\,x[\varphi]\,|\,x[\psi]\rangle = \langle x[\varphi]\,|\,D[a^*]^*\,|\,x[\psi]\rangle = \langle x[\varphi]\,|\,D[a]\,|\,x[\psi]\rangle, \tag{28}$$

and by continuity $D[a^*]^* = D[a]$ as was to be proved.

If a_1 and a_2 are two point transformations in $\mathfrak{J}$ satisfying the boundedness condition (17) it follows from (16) that

$$D[a_2]\,D[a_1] = D[a_2\,a_1]. \tag{29}$$

6. Symmetries of a stochastic process

A point transformation a on the parameter space of the stochastic process $x(.)$ will be called a symmetry transformation of the stochastic process if

$$\begin{aligned} &\text{a.} \quad r(a\,\tau_1, a\,\tau_2) = r(\tau_1, \tau_2) \quad (\text{arbitrary } \tau_1, \tau_2 \in \mathfrak{J}) \\ &\text{b.} \quad a \text{ has an inverse } a^{-1}, \text{ satisfying } a^{-1}a\,\tau = a\,a^{-1}\tau = \tau \quad (\text{arbitrary } \tau \in \mathfrak{J}). \end{aligned} \tag{30}$$

From (30) a. it is seen that the boundedness condition (17) is satisfied with $M_a = 1$, and we find

$$D[a]^* D[a] = 1, \tag{31}$$

that is $D[a]$ is an isometry of the Hilbert space $H(\mathfrak{J})$. From (30) b. we find that in (23) we may take $a^* = a^{-1}$ and using (29) we finally obtain

$$D[a]^{-1} = D[a]^* = D[a^{-1}], \tag{32}$$

that is, a symmetry operation of the stochastic process gives rise to a unitary transformation in the Hilbert space of the process.

The symmetries of a stochastic process, which evidently form a group, may be said to correspond to the integrals of motion of a dynamical system or rather to the transformations generated by the integrals of motion. The explicit representation of the stochastic process as a vector in the Hilbert space $H(\mathfrak{J})$ is greatly simplified if one uses a "coordinate system" in which as many of the symmetry transformations as possible are "diagonal". We see that there is a certain formal analogy between stochastic processes with symmetry operations and quantum mechanical systems with integrals of motion.

Let a be a symmetry operation of the stochastic process $x(.)$ and $\prod(\mathfrak{J}_1)$ the orthogonal projection onto the subspace $H(\mathfrak{J}_1)$ corresponding to the subset $\mathfrak{J}_1$ of the parameter set $\mathfrak{J}$. Then

$$\begin{aligned} D[a]\prod(\mathfrak{J}_1)D[a]^* &= \prod(a\,\mathfrak{J}_1) \quad \text{or} \\ D[a]\prod(\mathfrak{J}_1) &= \prod(a\,\mathfrak{J}_1)D[a]. \end{aligned} \tag{33}$$

To obtain the relation (33) we note that $D[a]\prod(\mathfrak{J}_1)D[a]^*$ is a projection, $D[a]$ being unitary. Take $x \in H(a\,\mathfrak{J}_1)$ that is $\prod(a\,\mathfrak{J}_1)x = x$. x is the limit of elements of the form $x[\psi]$, $\psi(.) \in G_0(a\,\mathfrak{J}_1)$. It follows that

$$D[a]\prod(\mathfrak{J}_1)D[a]^* x[\psi] = x[\psi],$$

and hence $D[a]\prod(\mathfrak{J}_1)D[a]^* x = x$. Conversely if $D[a]\prod(\mathfrak{J}_1)D[a]^* x = x$ we have $\prod(\mathfrak{J}_1)D[a]^* x = D[a]^* x$ and hence $\prod(a\,\mathfrak{J}_1)x = x$. It follows that the projections are equal.

7. Stationary vector processes

In this section we shall specialize the parameter set $\mathfrak{J}$ of the stochastic process $x(.)$. Each point in $\mathfrak{J}$ is assumed to be of the form

$$\tau = (t, u), \quad \mathfrak{J} = (R^1 \times V) \tag{34}$$

with the coordinate t taking values on the real axis and representing the time and the coordinate u being an element of the m-dimensional complex vector space V introduced in I, section 3. It is assumed that $x(t, u)$ depends linearly on u, that is

$$x(t,\, u_1 + u_2) = x(t,\, u_1) + x(t,\, u_2),\; x(t,\, \lambda u) = \lambda\, x(t,\, u). \tag{35}$$

It follows from (35) that the covariance function defines a linear operator $R(t_1, t_2)$ on V

$$\langle x(t_1,\, u_1) \,|\, x(t_2,\, u_2) \rangle = (u_1 \,|\, R(t_1,\, t_2) \,|\, u_2). \tag{36}$$

We shall call $R(t_1, t_2)$ the covariance operator. It has the property

$$R(t_1,\, t_2)^* = R(t_2,\, t_1) \tag{37}$$

and for arbitrary $t_1, \ldots, t_k$, $u_1, \ldots, u_k$

$$\sum_{r,s=1}^{r,s=k}\sum (u_r \,|\, R(t_r,\, t_s) \,|\, u_s) \geqslant 0 \tag{38}$$

corresponding to (7).

The relation (35) suggests the following notation

$$x(t,\, u) = (x(t) \,|\, u),\;\; \overline{x(t,\, u)} = (u \,|\, x(t)), \tag{39}$$

with $x(t)$ a random vector taking values in V and

$$R(t_1,\, t_2) = \mathbf{E}\,|\, x(t_1))\, (x(t_2) \,|. \tag{40}$$

The process $x(.,\, .)$ is assumed to have the following symmetries which we express as conditions on the covariance operator

$$\begin{aligned} &\text{a.}\qquad \text{Stationarity: } R(t_1,\, t_2) = R(t_2 - t_1),\\ &\text{b.}\qquad \text{Time reversal invariance: } T\,R(-t_1,\, -t_2)\,T^* = R(t_1,\, t_2), \end{aligned} \tag{41}$$

where T is the time reversal operator introduced in I, equation (3).

The unitary operators in $H(R^1 \times V)$ corresponding to (41) a. and b. are the group of translation operators $U(t)$

$$U(t_0)\, x(t,\, u) = x(t + t_0,\, u), \tag{42}$$

and the time reversal operator T

$$T\, x(t,\, u) = x(-t,\, T^* u). \tag{43}$$

We shall further assume that

$$\begin{aligned} &\text{a.}\qquad R(t) \text{ is continuous at the origin, that is } \lim_{t \to 0} R(t) = R(0)\\ &\text{b.}\qquad R(0) \text{ is non singular.} \end{aligned} \tag{44}$$

It follows from (44) a. that $U(t)$ is a (weakly) continuous function of t and that the Hilbert space of the process $H(R^1 \times V)$ is *separable* that is has a denumerable basis. (44) b. on the other hand is no essential restriction, for $R(0)u=0$ is equivalent to $x(0, u)=0$ and hence to $x(t, u)=U(t)x(0, u)=0$ for any value of t, and we may restrict ourselves to vectors u from the subspace of V on which $R(0)$ is non singular.

If the stochastic process $x(.\,,.)$ represents the values of an observable associated with a closed system described by classical mechanics or by quantum mechanics it is possible to derive the condition (41) b. from the mechanical laws. The translation operator $U(t)$ is in this case directly connected with the microscopic dynamics of the system and the time reversal invariance of the basic equations of motion is represented by the commutation relation

$$T\,U(t)\,T=U(-t) \tag{45}$$

and the unitarity of T.

8. Deterministic and nondeterministic processes. Dissipation

We introduce the simplified notation $H_t=H((-\infty, t]\times V)$. The subspace H_t consists of the full past of the stochastic process $x(.,.)$ at the time t and evidently H_{t+h} contains H_t for any positive h. It follows that the projection $\prod_t$ onto H_t is monotonically non decreasing as function of t. Hence the limits

$$\lim_{t\to-\infty}\prod_t=\prod_{-\infty} \text{ and } \lim_{t\to+\infty}\prod_t=\prod_{+\infty}{}^{1} \tag{46}$$

exist as welldefined projections on the Hilbert space of the process. It follows directly that $\prod_{+\infty}$ is the projection onto the Hilbert space of the process, for any element of that space can be approximated by an element of the form $x[\varphi]$ (5) and for any such element $\prod_{+\infty}x[\varphi]=x[\varphi]$. We observe that the limits (46) exist whether the process is stationary or not.

In general $\prod_{-\infty}\leqslant\prod_{+\infty}$. If $\prod_{-\infty}=\prod_{+\infty}$ we shall say that the stochastic process $x(.\,,.)$ is deterministic. The meaning of this condition is obviously that the future of the process at an arbitrary time t_1 is determined by the full past of the process at an arbitrary time t_2.

Now we take into explicit account the fact that the process is stationary. Applying the equality (33) to $\prod_t$ and $U(t)$ we find

$$U(t_1)\prod_{t_2}U(t_1)^*=\prod_{t_2+t_1} \tag{47}$$

and combining (46) and (47)

$$U(t)\prod_{-\infty}U(t)^*=\prod_{-\infty},\; U(t)\prod_{+\infty}U(t)^*=\prod_{+\infty}. \tag{48}$$

The relations (48) express the fact that the subspaces corresponding to the projections $\prod_{-\infty}$ and $\prod_{+\infty}$ are invariant under $U(t)$. It is easily seen that the

---

[1] The limits exist in the sense $\lim_{t\to{+\infty \atop -\infty}}||(\Pi_t-\Pi_{{+\infty \atop -\infty}})x||=0$ for any $x\in H(R^1\times V)$.

stochastic process is deterministic if and only if for two arbitrary fixed times t_1 and t_2, $t_1 < t_2$ we have

$$\prod_{t_1} = \prod_{t_2} \text{ or equivalently } H_{t_1} = H_{t_2}. \tag{49}$$

The "only if" proposition is trivially true. Conversely it is seen that (49) and the monotony of $\prod_t$ imply

$$\prod_{t_1} = \prod_{t_1+h} = U(h) \prod_{t_1} U(h)^* \quad (0 \leqslant h \leqslant t_2 - t_1) \tag{50}$$

and hence $\prod_{t_1} = U(h) \prod_{t_1} U(h)^* = U(2h) \prod_{t_1} U(2h)^* = \cdots = U(-h) \prod_{t_1} U(-h)^* = U(-2h) \prod_{t_1} U(-2h)^* = \cdots$. An arbitrary time t can be written in the form $t = t_1 + n_t \cdot h_t$ with n_t integer and $0 \leqslant h_t \leqslant t_2 - t_1$ and $\prod_t$ is found to be independent of t, so that $\prod_t = \prod_{-\infty} = \prod_{+\infty}$ and the process is deterministic.

If $\prod_{-\infty} = 0$ the stochastic process is said to be non deterministic. It is evident that in general the stochastic process $x(.,.)$ may be considered as a linear superposition of a deterministic part $\prod_{-\infty} x(.,.)$ and a non deterministic part $(\prod_{+\infty} - \prod_{-\infty}) x(.,.)$. Writing

$$x_0(t, u) = \prod_{-\infty} x(t, u) \text{ and } x_1(t, u) = x(t, u) - x_0(t, u) \tag{51}$$

for the deterministic and the non deterministic part respectively, we can write the covariance operator $R(t)$ as a sum of two covariance operators

$$R(t) = R_0(t) + R_1(t) \tag{52}$$

corresponding to the decomposition into two uncorrelated processes (51)

$$\begin{aligned} (u \,|\, R_0(t) \,|\, v) &= \langle x_0(t', u) \,|\, x_0(t'+t, v) \rangle \\ (u \,|\, R_1(t) \,|\, v) &= \langle x_1(t', u) \,|\, x_1(t'+t, v) \rangle. \end{aligned} \tag{53}$$

Evidently the stochastic process $x(.,.)$ is deterministic if and only if $R_1(0) = 0$ and non deterministic if and only if $R_0(0) = 0$. In general we have

$$\text{Rank}\{R_0(0) + R_1(0)\} = m, \; 0 \leqslant \text{Rank}\{R_0(0)\} \leqslant m, \; 0 \leqslant \text{Rank}\{R_1(0)\} \leqslant m. \tag{54}$$

It is seen that deterministic and non deterministic processes can be studied separately.

If the stochastic process $x(.,.)$ has no deterministic part and u is an arbitrary nonzero vector in V, it follows immediately that the (one-dimensional) stochastic process $x(., u)$ is non deterministic, for evidently the corresponding subspace $H((-\infty, t] \times \{u\})$ is contained in $H((-\infty, t] \times V)$ so that from (46):

$$\lim_{t \to -\infty} H((-\infty, t] \times \{u\}) = \{0\}, \tag{55}$$

that is the subspace consisting only of the null-vector, whereas as a consequence of the assumption $u \neq 0$

$$\lim_{t\to-\infty} H((-\infty, t]\times\{u\}) \tag{56}$$

contains at least one nonzero element. Hence it follows in the general case that if $u \in V$ is such that $R_1(0)\,u \neq 0$, then the (one dimensional) stochastic process $x_1(.,\,u)$ is non deterministic.

We shall study the non deterministic part of the stochastic process $x(.,.)$, assuming $\prod_{-\infty} \neq \prod_{+\infty}$. It is seen that F_t defined by

$$F_t = \prod_t - \prod_{-\infty} \tag{57}$$

is a spectral distribution[1] on the $U(t)$-invariant subspace corresponding to, the projection $\prod_{+\infty} - \prod_{-\infty}$, that is $H_{-\infty} \ominus H_{-\infty}$, the Hilbert space of the non deterministic component. Hence we may define the group of unitary transformations $V(\lambda)$ on $H_{+\infty} \ominus H_{-\infty}$ by

$$V(\lambda) = \int_{-\infty}^{\infty} \exp\{i\lambda t\}\, d\,F_t \quad (-\infty < \lambda < +\infty). \tag{58}$$

From (47), (48) and (51) we readily obtain the commutation relation

$$U(t)\,V(\lambda) = \exp\{-i\lambda t\}\,V(\lambda)\,U(t). \tag{59}$$

Here $U(t)$ is a continuous function of t according to our assumption (44) a. and $V(\lambda)$ is a continuous function of λ according to its definition (58).

We observe that (59) is an expression of the commutation relation for canonical coordinates and momenta in quantum mechanics. For writing heuristically

$$\begin{aligned} U(t) &= \exp\{i\,t\,q\} = 1 + i\,t\,q + \tfrac{1}{2}(i\,t\,q)^2 + \ldots \\ V(\lambda) &= \exp\{i\,\lambda\,p\} = 1 + i\,\lambda\,p + \tfrac{1}{2}(i\,\lambda\,p)^2 + \cdots \end{aligned} \tag{60}$$

we obtain by formal calculations from (59)

$$q\,p - p\,q = i\,1, \tag{61}$$

that is the Heisenberg form of the commutation relation. In this connection the equation (59) is sometimes called Weyl's relation [Weyl, 1927, and 1931 § 14, 15]. It is known that if q is represented by the multiplicative operator in the space $L^2(-\infty,\ +\infty;\ d\,q)$ the representation of p is determined to be

$$p = -i\frac{d}{d\,q} + \frac{d\,F(q)}{d\,q} = \exp\{-i\,F(q)\}\left[-i\frac{d}{d\,q}\right]\exp\{i\,F(q)\}, \tag{62}$$

where $F(q)$ is an "arbitrary" real function of q. That means that p is determined by q and (61) up to a unitary transformation commuting with q. A simple proof can be found in [Dirac, 1947] and a rigorous version in [von

---

[1] This concept is defined in Appendix 2.

Neumann, 1931]. For our purposes a somewhat more general representation theorem for the groups connected by (59) is needed. It is not convenient to work with the infinitesimal generators q and p which are unbounded operators in the Hilbert space of the stochastic process and hence cannot be defined for every element of that space. With the aid of this representation theorem, which is proved in appendix 2, it is possible to obtain an isometric connection between the Hilbert space of the stochastic process and the Hilbert space of m-dimensional vector functions on the real axis L^2_V, to be defined below, such that roughly speaking q is represented by the infinitesimal translation on the time axis $i\,d/dt$ and p is represented by the multiplicative operator, t. Before discussing this representation we shall be concerned with certain dissipative properties of non deterministic stationary stochastic processes.

The subspace $H(\{t\}\times V)$ corresponding to one single instant t is unitarily equivalent to the subspace $H(\{0\}\times V)$. We have assumed in section 7 that the operator $R(0)$ in V is non singular, and as was pointed out there is no essential restriction in this assumption. It follows that $x(0, u)$ may be considered as an invertible mapping of V on $H(\{0\}\times V)$

$$x(0, u) = X\,u; \quad u = X^{-1}\,x(0, u). \tag{63}$$

The adjoint of X, X^*, is a mapping of $H(\{0\}\times V)$ on V

$$\langle X\,u\,|\,x(0, v)\rangle = (u\,|\,X^*\,x(0, v)) = (u\,|\,X^*\,X\,|\,v) \;\text{ or }\; X^*\,X = R(0). \tag{64}$$

Let us now consider as in section 4 the (linear) estimate of $x(t, u)$ when the stochastic process has been observed up to the time 0. The best linear estimate is determined by the projection

$$\textstyle\prod_0 x(t, u) = \prod_0 U(t)\,x(0, u) = U(t)\prod_{-t} x(0, u), \tag{65}$$

and the estimation error $x(t, u) - \prod_0 x(t, u)$ has the variance

$$\begin{aligned}\textstyle\|x(t, u) - \prod_0 x(t, u)\|^2 &= \textstyle\|x(t, u)\|^2 - \|\prod_0 x(t, u)\|^2 = \\ &= \textstyle\|x(0, u)\|^2 - \|\prod_{-t} x(0, u)\|^2.\end{aligned} \tag{66}$$

It follows from the first of the relations (46) that the variance of the estimation error is a monotonic function of t, equal to zero for $t = 0$ and tending to its maximum value $\|x(0, u)\|^2 - \|\prod_{-\infty} x(0, u)\|^2$ when $t \to \infty$. If $x(.\,,.)$ is non deterministic this maximum value is equal to $\|x(0, u)\|$ otherwise it is smaller and in particular for a deterministic process it is equal to zero. It is suggestive to compare the monotonic increase of the estimation error variance with the monotonic increase of thermodynamic potentials and this point will be discussed further below.

From the last expression of the estimation error variance in (66) it can easily be seen that this quantity may be associated with an operator $\mathcal{E}(t)$ in the m-dimensional vector space V

$$(u\,|\,\mathcal{E}(t)\,|\,v) = \langle x(0, u)\,|\,x(0, v)\rangle - \textstyle\langle x(0, u)\,|\prod_{-t}|\,x(0, v)\rangle \tag{67}$$

or $\mathcal{E}(t)=R(0)-X^*\prod_{-t}X$, where X is the mapping defined by (64). The operator $\mathcal{E}(t)$ is non negative, $(u|\mathcal{E}(t)|u)\geqslant 0$ for arbitrary u in V, and increases monotonically from $\mathcal{E}(0)=0$ to a maximal value $\mathcal{E}(\infty)$:

$$\mathcal{E}(\infty)=R(0)-X^*\prod_{-\infty}X. \tag{68}$$

It may be noted that the derivative of $\mathcal{E}(t)$ exists and is non negative. This is a consequence of the fact that, as proved in appendix 2, the "spectral distribution" $\prod_t$ is absolutely continuous. The stochastic process $x(.,.)$ has a non deterministic component if and only if for some t, e.g. $t=0$, we have

$$\frac{d}{dt}\mathcal{E}(t)\neq 0. \tag{69}$$

The derivative is necessarily non negative.

It must be noticed that the estimation discussed above is built on the observation of the full past of the stochastic process. If the full past of the process is not known the estimation has to be modified in an obvious way. In this case we can in general not expect a monotonic behaviour of the estimation error variance. We consider the extreme case of estimating $x(t, u)$ from a single observation which gives us $x(0, .)$. The estimate $x(t, u, \{0\})$ is determined by the projection of $x(t, u)$ into $H(\{0\}\times V)$, that is by the scalar products

$$\begin{gathered}\langle x(0, v)|x(t, u, \{0\})\rangle=\langle x(0, v)|U(t)|x(0, u)\rangle\\ (v|R(0)|X^{-1}x(t, u, \{0\}))=(v|R(t)|u),\end{gathered} \tag{70}$$

and hence

$$\begin{gathered}X^{-1}x(t, u)\{0\})=R(0)^{-1}R(t)u \text{ or}\\ x(t, u, \{0\})=x(0, R(0)^{-1}R(t)u).\end{gathered} \tag{71}$$

The estimation error variance is given by

$$(u|R(0)-R(t)^*R(0)^{-1}R(t)|u). \tag{72}$$

It follows that the operator

$$R(0)-R(t)^*R(0)^{-1}R(t)\geqslant 0 \tag{73}$$

is equal to zero for $t=0$, but in general its dependence on t is not monotonic. If the m-dimensional Gauszian stochastic process corresponding to the covariance operator $R(t)$ (see below, section 11) is mixing, that is if

$$\lim_{t\to\infty} R(t)=0 \tag{74}$$

we find that the estimation error variance approaches its maximal value $R(0)$ when t tends to infinity.

9. Causal functions and transforms

We have seen in the last section that a non deterministic stochastic process is always associated with the two groups of unitary transformations $U(.)$ and $V(.)$ which are related by the commutation relation (59). As we shall see in the next section the existence of these two groups is sufficient to establish a very simple representation of the stochastic process. The framework for this representation is the Hilbert space of m-dimensional vector functions on the real axis, L^2_V, and this section is devoted to a study of this space.

Let V be the m-dimensional complex vector space introduced in section 7. By L^2_V we denote the class of all functions on the real axis with values in V, such that

$$\int_{-\infty}^{+\infty} (v_t | v_t)\, dt < +\infty . \tag{75}$$

It is easily shown that L^2_V is a Hilbert space with the scalar product

$$\langle v. | u. \rangle = \int_{\infty}^{+\infty} (v_t | u_t)\, dt \tag{76}$$

and the correspoding norm $\| v. \|^2 = \langle v. | v. \rangle$.

On L^2_V we introduce the group of translations $U_z(.)$

$$[U_z(t)\, v.]_{t'} = v_{t'-t} \qquad (v. \in L^2_V, \ -\infty < t < +\infty), \tag{77}$$

and the multiplicative group $V_z(.)$

$$[V_z(t)\, v.]_{t'} = \exp\{i\, t\, t'\}\, v_{t'}. \tag{78}$$

It follows at once from the definitions of $U_z(.)$, $V_z(.)$ and the scalar product (76) that the transformations (77) and (78) are unitary and that they are connected by the commutation relation

$$U_z(t)\, V_z(t')\, U_z(t)^{-1}\, V_z(t')^{-1} = \exp\{-i\, t\, t'\}, \tag{79}$$

which is just the Weyl relation (59).

We shall further introduce the transformation J, inversion at the origin,

$$[J\, v.]_t = v_{-t}. \tag{80}$$

The inversion transformation is obviously a unitary involution

$$J^2 = 1, \ J^* = J^{-1} = J, \tag{81}$$

and connected with the group of translations by the "time reversal relation"

$$J\, U_z(t) = U_z(-t)\, J, \tag{82}$$

$$J\, V_z(t) = V_z(-t)\, J.$$

Complex conjugation may be represented as an anti linear, isometric transformation C on L_V^2

$$[C\,v.]_t = C\,v_t \tag{84}$$

where $C\,v_t$ denotes the complex conjugation operation in V, as defined in I, section 3.

From the definitions $U_z(.)$ and $V_z(.)$ it follows at once that

$$\begin{aligned} C\,U_z(t) &= U_z(t)\,C, \\ C\,V_z(t) &= V_z(-t)\,C. \end{aligned} \tag{85}$$

The Fourier-Plancherel transformation S in L_V^2 is defined by

$$[S\,v.]_t = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\{-i\,t\,t'\}\,d\,t'\,v_{t'}, \tag{86}$$

where the integral is defined as a limit in the mean or as

$$[S\,v.]_t = \frac{1}{\sqrt{2\pi}} \frac{d}{d\,t} \int_{-\infty}^{\infty} \frac{\exp\{-i\,t\,t'\}-1}{-i\,t'}\,d\,t'\,v_{t'}. \tag{87}$$

The Fourier-Plancherel transformation S is unitary (Parseval's theorem) and satisfies

$$S^2 = J, \quad S^4 = 1, \tag{88}$$

and hence
$$S^* = S\,J = J\,S = S^3 = S^{-1}. \tag{89}$$

Further it is obvious that
$$C\,S = S^*\,C. \tag{90}$$

By direct application of the definitions it is found that $U(.)$ and $V(.)$ are connected by the Fourier-Plancerel transformation

$$U_z(t)\,S = S\,V_z(t), \tag{91}$$

or
$$S^*\,U_z(t)\,S = V_z(t); \quad S^*\,V_z(t)\,S = J\,U_z(t)\,J = U_z(-t). \tag{92}$$

A family of projection operators $\{F_{z,t} \mid -\infty < t < +\infty\}$ is defined by

$$[F_{z,t}\,v.]_{t'} = \begin{matrix} v_{t'} & t' \leqslant t \\ 0 & t' > t. \end{matrix} \tag{93}$$

Evidently $F_{z,.}$, when properly normalized, is a spectral distribution corresponding to the spectral resolution of the group $V_z(.)$

$$V_z(t) = \int_{-\infty}^{+\infty} \exp\{i\,t\,t'\}\,d_{t'}\,F_{z,t'}. \tag{94}$$

A vectorfunction $v. \in L_V^2$ is called *causal* if

$$F_{z,0}\, v. = 0, \tag{95}$$

and a vectorfunction $u. \in L_V^2$ is called a *causal transform* if its Fourier transform is a causal function, that is if

$$F_{z,0}\, S\, u. = 0. \tag{96}$$

The concepts causal function and transform have arisen in connection with the study of so called stationary causal filters, which are represented by linear, translation invariant transformations which have the property of transforming causal functions into causal functions, and are characterized by "dispersion relations"; see e.g. [Toll, 1956].

The projection onto the subspace of causal transforms is according to (96)

$$P = S^*\, F_{z,0}\, S. \tag{97}$$

It follows from (90), (89), (92) that P is connected with the conjugation through

$$C\,P = (1 - P)\, C. \tag{98}$$

We introduce the additive but non homogeneous operations "real part" and "imaginary part"

$$\mathcal{R} = \tfrac{1}{2}\,(1 + C), \tag{99}$$

$$\mathcal{J} = \frac{1}{2\,i}\,(1 + C), \tag{100}$$

and the real transformation $\mathcal{H}$ in L_V^2 (the Hilbert transformation)

$$\mathcal{H} = 2\,\mathcal{J}\,P\,\mathcal{R} = i\,(1 - 2\,P)\,\mathcal{R}, \tag{101}$$

$$\mathcal{R}\,\mathcal{H} = \mathcal{H}\,\mathcal{R} = \mathcal{H}. \tag{102}$$

It follows at once from the second expression (101) that

$$\mathcal{H}\,\mathcal{H} = -\,\mathcal{R} \tag{103}$$

and that
$$\|\,\mathcal{H}\,v.\| = \|\,\mathcal{R}\,v.\| \qquad (v. \in L_V^2), \tag{104}$$

Hence $\mathcal{H}$ is an isometric transformation of the *real* Hilbert space $\mathcal{R}\,L_V^2$ onto itself, and its inverse is $-\,\mathcal{H}$. It follows further that with $v.$ arbitrary in L_V^2, or $\mathcal{R}\,v.$ arbitrary in $\mathcal{R}\,L_V^2$

$$\mathcal{R}\,v. + i\,\mathcal{H}\,\mathcal{R}\,v. = 2\,P\,\mathcal{R}\,v. \tag{105}$$

is always a causal transform. On the other hand we have

$$\mathcal{H}\mathcal{R}P=\mathcal{J}P \qquad \text{or for } Pv.=v. \qquad \mathcal{H}\mathcal{R}v.=\mathcal{J}v. \tag{106}$$

$$\mathcal{H}\mathcal{J}P=-\mathcal{R}P, \qquad \text{or for } Pv.=v. \qquad \mathcal{H}\mathcal{J}v.=-\mathcal{R}v. \tag{107}$$

showing that a causal transform is always of the form (105). We have obtained by algebraic methods the wellknown criterion [Titchmarsh, 1948]: a function is a causal transform if and only if its real and imaginary parts are connected by a Hilbert transformation. Usually the criterion is derived from the explicit representation of the Hilbert transform

$$[\mathcal{H}\mathcal{R}v.]_t=\frac{1}{\pi}\ \text{Princ. value} \int_{-\infty}^{+\infty} \frac{1}{s-t}\mathcal{R}v_s\,ds, \tag{108}$$

which is easily obtained from (101).

There is a further wellknown criterion: a function is a causal transform if and only if it can be extended to a function which is analytic in the upper half of the complex plane (for the precise formulation see e.g. [Titchmarsh, 1948]).

So far the fact that we are dealing with vector valued functions has played no essential role as all statements can be referred to complex functions of the form $(u|v.)$ with $u\in V$ and $v.\in L_V^2$.

We shall now consider a family of linear transformations $f(\lambda)$ on the m-diminsional vector space V, defined for every real λ, such that for every vector $v\in V$ the vector function $J f(\lambda)\,v$ is a causal transform. This implies in particular that for every v, $f(.)\,v$ is an element of L_V^2, and hence the integral

$$\int_{-\infty}^{+\infty} f(\lambda)^* f(\lambda)\,d\lambda \tag{109}$$

exists and defines a non negative operator in V. Further the Fourier-Plancherel transform of $f(.)\,v$

$$Sf(.)\,v=g(.)\,v \tag{110}$$

defines a family of linear transformations $g(.)$ on V with the property

$$\begin{aligned} &F_{z,0}\,g(.)\,v=g(.)\,v \\ \text{or} \qquad &g(t)=0 \qquad (t>0). \end{aligned} \tag{111}$$

From the unitarity of the Fourier-Plancherel transformation S and the relation (92) it follows in view of (111) that

$$\int_{-\infty}^{+\infty} \exp\{i\lambda t\}\,f(\lambda)^* f(\lambda)\,d\lambda=\int_{-\infty}^{\min(0,t)} g(t')^*\,g(t'-t)\,dt'. \tag{112}$$

It is not difficult to see that one can always find families of transformations $f(.)$ with the desired properties. Excluding the trivial case $f(.)=0$, we can always find a finite real number $t_0=0$ such that t_0 is the smallest number for which

$$\int_{\infty}^{0} g(t')^{*} g(t')\,dt' = \int_{-\infty}^{t_0} g(t')^{*} g(t')\,dt'. \tag{113}$$

It is no restriction to assume $t_0 = 0$, for if this is not orginally the case we can replace $f(\lambda)$ by $\exp\{i\lambda t_0\} f(\lambda)$.

We consider now the subspace of L_V^2 which is generated by all elements of the form

$$\sum_{t \leqslant 0} \sum_{u \in V} c(t, u)\, U_z(t)\, g(.)\, u \tag{114}$$

where the complex function $c(t, u)$ is equal to zero for all but a finite number of values of its arguments. Evidently this subspace which we shall denote by $L(0)$ is contained in the subspace $F_{z,0} L_V^2$ formed by all vector functions in L_V^2 which vanish on the positive part of the real axis. In general $L(0)$ will be a proper subspace of L_V^2. The family of subspaces $L(t)$ defined by

$$L(t) = U(t)\, L(0), \tag{115}$$

is evidently increasing with t: $L(t+h)$ contains $L(t)$ if $h \geqslant 0$. Hence the limits $L(-\infty)$ and $L(+\infty)$ exist, and evidently

$L(-\infty) = \{0\}$ (the subspace consisting only of the 0-vector)

$L(+\infty) \subset L_V^2$.

It follows from the form of the elements (114) that $L(+\infty)$ consists exactly of all vector functions of the form

$$f(\lambda)\, v_\lambda, \quad \int_{-\infty}^{+\infty} (v_\lambda | f^{*}(\lambda) f(\lambda) | v_\lambda)\, d\lambda < +\infty. \tag{116}$$

Hence we find that $L(+\infty) = L_V^2$ if and only if $f(\lambda)$ has an inverse $f(\lambda)^{-1}$ for every value λ. In this case only every element of L_V^2 can be represented in the form (116).

It follows further that $L(+\infty) = L_V^2$ if and only if $L(0) = F_{z,0} L_V^2$. Evidently the condition is sufficient for $L(0) = F_{z,0} L_V^2$ implies that $L(t) = F_{z,t} L_V^2$ for every value t. To see that the condition is also necessary we assume that $L(+\infty) = L_V^2$. As we have seen this is equivalent to the existence of $f(\lambda)^{-1}$ for every λ. The elements of $L(0)$ are the Fourier-Plancherel transforms of elements of the form $f(\lambda)\, v_\lambda$ such that $J f(.) v.$ is a causal transform. On the other hand as $f(\lambda)^{-1}$ exists every causal transform can be represented in this form, and the proposition is proved.

We shall conclude this section with a discussion of the following problem. Given a family of non singular positive transformations $\sigma'(\lambda)$ on V defined for every real λ. It is assumed that $\sigma'(.)$ is integrable over the real axis. Under which conditions is it possible to represent $\sigma'(.)$ in the form

$$\sigma'(\lambda) = f(\lambda)^{*} f(\lambda) \tag{117}$$

such that $f(\lambda)v$ is a causal transform for every $v \in V$. Evidently a necessary and sufficient condition is that the Fourier-Plancherel transform of $\sigma'(\lambda)$ can be represented in the form (112)

$$\int_{-\infty}^{+\infty} \sigma'(\lambda) \exp\{i\lambda t\}\, d\lambda = \int_{\infty}^{\min(0,t)} g(t')^* g(t'-t)\, dt'. \tag{118}$$

It is known that in the one-dimensional case ($\dim V = m = 1$) a necessary and sufficient condition for the possibility of a factoring (117) is that

$$\int_{\infty}^{-\infty} \frac{\log \sigma'(\lambda)}{1+\lambda^2}\, d\lambda > -\infty \tag{119}$$

and that the factor $f(\lambda)$ is given by

$$f(\lambda) = \exp\left\{-\frac{1}{2\pi i}\int_{-\infty}^{+\infty} \frac{1+t\lambda}{t-\lambda}\, \frac{\log \sigma'(t)}{1+t^2}\, dt\right\}. \tag{120}$$

The factoring problem is identical with the problem of determining the scattering phase shift from the scattering amplitude [Toll, 1956]. In [Karhunen, 1950] the general solution of the factoring problem is derived.

One can expect that in the multipdimensional case (119) and (120) hold as operator relations in V. However due to computational difficulties such relations are almost useless. On the other hand it is easy to derive some conditions which are only necessary. E.g.: the factoring problem (117) has a solution only if for each non zero element v in V

$$\int_{\infty}^{+\infty} \frac{\log (v|\sigma'(\lambda)|v)}{1+\lambda^2}\, d\lambda > -\infty \tag{121}$$

or, the factoring problem has a solution only if

$$\int_{-\infty}^{+\infty} \frac{\log\{\mathrm{Trace}\ \sigma'(\lambda)\}}{1+\lambda^2}\, d\lambda > -\infty. \tag{122}$$

10. Representation of a non deterministic process

We shall now return to the stochastic process $x(.,.)$ introduced in section 7. According to Stone's theorem on the representation of one-parametric groups of unitary transformations, the translation operator $U(t)$ has the spectral resolution (see appendix 2)

$$U(t)=\int_{-\infty}^{+\infty} \exp\{it\lambda\}\, dE_\lambda, \tag{123}$$

corresponding to the Cramér-representation of the stochastic process

$$x(t,v)=\int_{-\infty}^{+\infty} \exp\{it\lambda\}\, d_\lambda z(\lambda,v), \tag{124}$$

where $z(.,.)$ is a stochastic process with uncorrelated increments

$$z(\lambda,v)=(E_\lambda-E_0)\,x(0,v), \tag{125}$$

and corresponding to the Cramér-Khinchin representation of the covariance operator $R(t)$ (section 7)

$$R(t)=\int_{-\infty}^{+\infty} \exp\{it\lambda\}\, d\sigma(\lambda) \tag{126}$$

where the spectral distribution operator on V, $\sigma(.)$, is determined by

$$(u\,|\,\sigma(\lambda)\,|\,v)=\langle x(0,u)\,|\,E_\lambda\,|\,x(0,v)\rangle. \tag{127}$$

It follows that

$$\langle d_{\lambda_1} z(\lambda_1,v)\,|\,d_{\lambda_2} z(\lambda_2,u)\rangle=\delta_{\lambda_1,\lambda_2}\, d_{\lambda_1}(v\,|\,\sigma(\lambda_1)\,|\,u) \tag{128}$$

where $\delta_{\lambda_1,\lambda_2}$ is Kronecker's symbol.

Applying the decomposition theorem proved in appendix 1 to the spectral distribution E, we obtain a decomposition of the stochastic process $x(.,.)$ into three mutually uncorrelated processes having respectively an absolutely continuous, a singular discrete (jump), and a singular continuous spectral distribution operator on V

$$\begin{aligned}
&\sigma(\lambda)=\sigma^a(\lambda)+\sigma^{sd}(\lambda)+\sigma^{sc}(\lambda)\\
&\sigma^a(\lambda)=\int_{-\infty}^{\lambda}\sigma'(\mu)\,d\mu,\quad (u\,|\,\sigma^a(\lambda)\,|\,v)=\langle x(0,u)\,|\,E^a_\lambda\,|\,x(0,v)\rangle\\
&\sigma^{sd}(\lambda)=\sum_{\lambda_\nu\leqq\lambda}\sigma(\lambda_\nu+0)-\sigma(\lambda_\nu-0),\quad (u\,|\,\sigma^{sd}(\lambda)\,|\,v)=\langle x(0,u)\,|\,E^{sd}_\lambda\,|\,x(0,v)\rangle\\
&\sigma^{sc}(\lambda)\colon\ (u\,|\,\sigma^{sc}(\lambda)\,|\,v)=\langle x(0,u)\,|\,E^{sc}_\lambda\,|\,x(0,v)\rangle.
\end{aligned} \tag{129}$$

It is easily seen that this decomposition of the stochastic process is invariant under the time reversal transformation T introduced in section 7 eq. (43) so that each of the three components of the process satisfies the symmetry relation (41 b), (see also [Uhlhorn, 1960 b]). The decomposition of the stochastic process into a deterministic and a non deterministic part was described in section 8.

It was shown in this section that the non-deterministic part of the stochastic process is characterized by the Weyl relation (59). On the other hand it is demonstrated in appendix 2 section 4. that the Weyl relation implies that the spectral distribution $E.$ is absolutely continuous. Hence the singular parts of the spectral distribution operator (129) correspond to deterministic stochastic processes whereas the absolutely continuous part $\sigma^a(.)$ may in general determine both a deterministic and a non deterministic stochastic process. In accordance with equations (52), (53)

$$\sigma(\lambda)=\sigma_0(\lambda)+\sigma_1(\lambda);\ \ \sigma^a(\lambda)=\sigma_1(\lambda)+\sigma_0^a(\lambda), \tag{130}$$

$$x(t, v)=x_{a,0}(t, v)+x_{sd}(t, v)+x_{sc}(t, v)+x_1(t, v). \tag{131}$$

We are now going to study the non deterministic component $x_1(.,.)$ of the stochastic process, assuming rank $R_1(0)=0$. In view of the orthogonal decomposition (131) we may avoid clumsy notations by assuming for the moment that the stochastic process consists only of a non deterministic component $x(\ldots)=x_1(\ldots)$.

The Hilbert subspace $H_t=H((-\infty, t]\times V)$ of the process (section 8) is generted by all finite sums of the type

$$\begin{aligned}\sum_{t'\leqslant t,}\sum_{u\in V} c(t', u)\,x(t', u)&=\sum_{t'\leqslant t,}\sum_{u\in V} c(t', u)\,U(t')\,x(0, u)=\\&=\sum_{t'\leqslant t,}\sum_{u\in V} c(t', u)\int_{-\infty}^{+\infty}\exp\{i t'\lambda\}\,d_\lambda z(\lambda, u).\end{aligned} \tag{132}$$

As $\sigma(.)$ is absolutely continuous and non negative we can always write

$$\sigma(\lambda)=\int_{-\infty}^{\lambda} f(\mu)^*\,f(\mu)\,d\mu. \tag{133}$$

Combining (132) and (133) we find that every element in H_∞ is uniquely mapped onto an element of the Hilbert space of m-dimensional vectorfunctions L_V^2 introduced in section 9. The mapping is linear and isometric and is determined by the correspondence

$$\begin{aligned}x(0, v)&\rightarrow f(\lambda)\,v\\U(t)\,x(0, v)&\rightarrow\exp\{i t\lambda\}\,f(\lambda)\,v.\end{aligned} \tag{134}$$

As is seen from the discussion of section 9 this mapping, determined by (134), is invertible exactly if the operator $f(\lambda)$ has an inverse for every λ. This condition is indeed satisfied for if v is a nonzero element in V we have for every λ

$$(v\,|\,\sigma(\lambda+h)-\sigma(\lambda)\,|\,v)>0\qquad(h>0), \tag{135}$$

and hence the operator $\sigma(\lambda+h)-\sigma(\lambda)$ has rank m if $h\neq 0$. The inequality (135) is a consequence of the fact that for fixed $v\in V$, $x(., v)$ is a one-dimensional non deterministic stochastic process (section 8) and the fact that the spectral distribution corresponding to a non deterministic process can have no gaps (appendix 2, section 5). From the connection between $\sigma(.)$ and $f(.)$ (133) it follows that $f(\lambda)^{-1}$ exists for almost every λ. We shall denote the unique inverse of the mapping (134) by Y_1. That is Y_1 is a linear isometric transformation of the Hilbert space L_V^2 onto the Hilbert space H_∞ of the stochastic process:

$$\begin{aligned} Y_1 f(.)\, v &= x(0, v) \\ Y_1^{-1} U(t)\, Y_1 &= V_z(t) \qquad \text{(eq. (78)).} \end{aligned} \tag{136}$$

Hence the L_V^2-representative of $U(t)$ is the multiplicative transformation $V_z(t)$. We want to determine the representative of $V(t)$ (58), $Y_1^{-1} V(t)\, Y_1$. It follows from the representation theorem for the Weyl relation proved in appendix 2 that $Y_1^{-1} V(t)\, Y_1$ is equivalent to the translation operator $U_z(-t)$ (77) up to a unitary transformation W which commutes with $Y_1^{-1} U(.)\, Y_1$, that is

$$\begin{aligned} U_z(-t) &= W^*\, Y_1^{-1}\, V(t)\, Y_1\, W \\ V_z(t) &= W^*\, Y_1^{-1}\, U(t)\, Y_1\, W = Y_1^{-1}\, U(t)\, Y_1. \end{aligned} \tag{137}$$

We define the isometric, linear transformations Y and Z of L_V^2 onto H_∞ by

$$Y = Y_1 W, \tag{138}$$

$$Z = Y S^* \quad (S \text{ defined by (86)}). \tag{139}$$

Combining (137), (138), (139), and (91) we find

$$U(t)\, Z = Z\, U_z(t) \qquad \text{(eq. (77)),} \tag{140}$$

$$V(t)\, Z = Z\, V_z(t) \qquad \text{(eq. (78)).} \tag{141}$$

From the definition of $V(t)$ (58) and of $V_z(t)$ (78), (94) it follows the correspondence between the spectral distributions $F_.$ and $F_{z,.}$

$$Z^{-1} F_t\, Z = F_{z,t}. \tag{142}$$

The transformation $Y^{-1} = (Y_1 W)^{-1} = W^* Y_1^{-1}$ can be analytically characterized by a correspondence of the form (134). This follows from the fact that (see appendix 2) the general form of a unitary operator in L_V^2, commuting with $V_z(.)$ is

$$[W v.]_\lambda = w(\lambda)\, v_\lambda \tag{143}$$

where $w(\lambda)$ is a unitary operator in the m-dimensional space V defined for every real λ. Hence, replacing in (134) $f(\lambda)$ by $w(\lambda)^* f(\lambda)$, we obtain the representation of Y^{-1}. In the following we shall denote the representative of Y^{-1} by $f(\lambda)$:

$$Y^{-1} x(0, v) = f(.)\, v. \tag{144}$$

Further we introduce in correspondence to eq. (110)

$$Z^{-1} x(0, v) = S f(.)\, v = g(.)\, v. \tag{145}$$

We have obtained the following representations of the non deterministic stochastic process $x(.,.)$

$$x(t, v) = Y V_z(t) f(.)\, v = Z U_z(t) g(.)\, v. \tag{146}$$

Here $J f(.)\, v$ is a causal transform for every $v \in V$. This follows from (142) and the definition of F_t. As

$$F_{z,0}\, g(.)\, v = F_{z,0}\, Z^{-1} x(0, v) = Z^{-1} F_0\, x(0, v) = Z^{-1} x(0, v) = g(.)\, v, \tag{147}$$

we have
$$F_{z,0}\, S J f(.)\, v = F_{z,0}\, J g(.)\, v = 0,$$

and the statement is proved.

Corresponding to (145) we obtain a representation of the covariance operator $R(t)$ (36), (118)

$$R(t) = \int_{-\infty}^{+\infty} f(\lambda)^* f(\lambda) \exp\{i \lambda t\}\, d\lambda = \int_{-\infty}^{\min(0,\, t)} g(t')^* g(t'-t)\, dt'. \tag{148}$$

We summarize the properties of the transformation Z from L_V^2 onto H_∞.

$$\begin{aligned} &\text{a. } Z \text{ is an invertible, isometric linear transformation.} \\ &\text{b. } U(t) Z = Z U_z(t). \\ &\text{c. } V(t) Z = Z V_z(t) \quad \text{or equivalently } F_0 Z = Z F_{z,0}. \end{aligned} \tag{149}$$

The transformation Z is uniquely determined by $g(.)$ (145) or equivalently by $f(.)$. Conversely $g(.)$ and $f(.)$ are uniquely determined by the properties (149) up to a unitary transformation in the m-dimensional space V. To see this, assume that Z_1 is a transformation satisfying (149). It follows directly that $Z_1^{-1} Z$ is a unitary transformation on L_V^2 which commutes with $U(.)$ and with $V(.)$. Hence $Z_1^{-1} Z$ is determined by a unitary operator in V, say w and

$$w g(.)\, v = Z_1^{-1} Z g(.)\, v = Z_1^{-1} x(0, v) = g_1(.)\, v \tag{150}$$

for every $v \in V$.

The transformation Z may be represented as a stochastic process with uncorrelated increments, that is a "random measure". Let $\varphi(I_a^b, .)$ be the indicator function

$$\varphi(I_a^b, t) = \begin{matrix} 1 & (a < t \leqslant b) \\ 0 & (t \leqslant a,\ t > b). \end{matrix} \tag{151}$$

For every $v \in V$ we have $\varphi(I_a^b, .)\, v \in L_V^2$. Hence we may define $Z(I_a^b, .)$ by

$$Z(I_a^b, v) = Z \varphi(I_a^b, .)\, v. \tag{152}$$

It follows at once from (149) that

$$\begin{aligned} & Z(I_a^b, \lambda u + \mu v) = \lambda Z(I_a^b, u) + \mu Z(I_a^b, v) \\ & \langle Z(I_a^b, u) \,|\, Z(I_a^b, v)\rangle = |b-a| (u|v) \\ & \langle Z(I_a^b, u) \,|\, Z(I_{a'}^{b'}, v)\rangle = 0 \qquad (b \leqslant a') \\ & U(t)\, Z(I_a^b, u) = Z(I_{a+t}^{b+t}, u) \\ & \langle x(t, u) \,|\, Z(I_a^b, v)\rangle = 0 \qquad (t \leqslant a), \end{aligned} \tag{153}$$

and that the representation (146) may be written in the symbolic form of a "causal" stochastic integral

$$x(t, v) = \int_{-\infty}^{t} Z(dt', g(t'-t)\, v). \tag{154}$$

In the one-dimensional case the representation (154) was obtained already in 1950 [Hanner, 1950], [Karhunen, 1950]. The methods of these papers are however not easily extended to the multidimensional case discussed here.

11. Application of the linear theory of stochastic processes to statistical mechanics

Before taking up the discussion of problems connected with the probabilistic interpretation of the linear theory of irreversible processes, outlined in I we shall make some general comments on the application of the linear theory of stochastic processes to the statistical description of mechanical systems of macroscopic dimensions. The stochastic processes which are used for the temporal description of macroscopic phenomena may always be associated with *stationary* processes. This possibility has been discussed in [Uhlhorn, 1960, b], and will be accepted in the present discussion. The stationary processes in question correspond to closed mechanical systems in macroscopic equilibrium. The stationary probability distributions which describe the fluctuations of the macroscopic observables in the equilibrium state are usually assumed to be sufficiently characterized by their first and second order moments: the mean values and the covariances of the observables. The justification of this assumption rests more or less directly on the law of large numbers, in particular on the central limit theorem. Formally the assumption is expressed by the statement that the probablility distribution in question are approximately Gauszian. By the same argument one can claim that the probability distributions of values observed at different times are approximately Gauszian, though the approximation may be worse than for the stationary distribution. In principle deviations from the Gauszian form can be obtained in a systematic way by the use of a multidimensional Cramér-expansion in terms of a parameter proportional to the size of the mechanical system [Uhlhorn, 1960 a]. One may even approximately replace the stationary stochastic process which describes the kinetics of fluctuations by a Gauszian process having the same first and second order moments.

Let us assume that the stationary stochastic process $x(.,.)$ introduced in section 7. corresponds to the measurement of an m-tuple of macroscopic observables. We assume that the mean values are normalized to zero and that the observables are real

$$\mathbf{E}\, x(t, u) = 0.$$

The probability distribution of any finite set $x(t_1, u_1), \ldots, x(t_p, u_p)$ can be obtained from the characteristic function, which in the Gauszian approximation is equal to

$$\mathbf{E} \exp \left\{ i \sum_{j=1}^{j=p} k_j\, x(t_j, u_j) \right\} = \exp \left\{ -\frac{1}{2} \sum_{i,j=1}^{i,j=p}\sum k_i\, k_j\, (u_i \,|\, R(t_j - t_i) \,|\, u_j) \right\} \quad (k_1, \ldots, k_p \text{ real}). \tag{155}$$

The family of all finite dimensional characteristic functions characterize the stochastic process. In the Gauszian approximation the covariance operator $R(t)$ is sufficient to characterize the process. We recall that Gauszian random variables are stochastically independent if and only if they are uncorrelated, corresponding to a factoring of the characteristic function.

From the form (155) of the characteristic function it is also easy to obtain the theorem, proved by Grenander in the one-dimensional case [Grenander, 1950]: A stationary, Gauszian stochastic process whose covariance operator $R(t)$ has the spectral representation

$$R(t) = \int_{\infty}^{\infty} \exp\{i \lambda t\}\, d\sigma(\lambda) \tag{156}$$

is ergodic exactly if the spectral distribution operator $\sigma(\lambda)$ is a continuous function of λ.

It is obvious from the Cramér-representation (124) of the Gauszian process that the condition is necessary. For if $\sigma(\lambda)$ has a jump at $\lambda \quad \lambda_0$ there exists a $v_0 \in V$ such that $z(\lambda + 0, v_0) - z(\lambda - 0, v_0) = z_{\lambda_0}$ is a non-vanishing complex Gauszian random variable with mean zero, on the probability space of the stochastic process, satisfying

$$U(t)\, z_{\lambda_0} = \exp(i t \lambda_0)\, z_{\lambda_0},$$

i.e. $U(t)$ has an eigenvector. The stochastic variable $|z_{\lambda_0}|$ is obviously $U(t)$-invariant but not equal to a constant. Hence the stochastic process is not ergodic if $\sigma(\lambda)$ is not continuous. That the condition is also sufficient is most easily proved by the same method as in [Uhlhorn, 1960 b] for the case of discrete time.

Usually we have no physical reasons to assume that $\sigma(\lambda)$ has a singular continuous component (though such distributions can easily be constructed). Hence if the Gauszian process is ergodic its covariance operator will usually have the property

$$\lim_{|t| \to \infty} R(t) = \lim_{|t| \to \infty} \int_{-\infty}^{+\infty} \exp\{i t \lambda\}\, \sigma'(\lambda)\, d\lambda = 0. \tag{157}$$

On the other hand (157) is sufficient to assure that the stochastic process is ergodic, and it in fact implies strong mixing as defined in [Hopf, 1937] and corresponding to complete stirring in Gibbs' sense [Gibbs, 1902, Ch. 12] [Uhlhorn, 1960 b]. In particular (157) implies that the conditional meanvalue of the m-dimensional observable described by $x(\ldots)$ will approach its equilibrium value in the course of time: If the value $v \in V$ was observed at $t=0$, the conditional meanvalue at an arbitrary time t is according to (71)

$$v(t) = R(t)^{*} R(0)^{-1} v; \;\; v(0) = v, \tag{158}$$

and hence

$$\lim_{t \to +\infty} v(t) = 0. \tag{159}$$

It should be remarked that the fluctuation around the conditional mean value (158) is determined by (72) and is always smaller than in the stationary state. Hence the meanvalue is representative for the process.

If the stationary Gauszian process $x(.,.)$ is non deterministic its spectral distribution operator $\sigma(\lambda)$ is certainly absolutely continuous, and the process is strongly mixing.

We remark that the time average of the stochastic process $x(t, u)$ is equal to its phase average (= mean value)

$$\lim_{t_0 \to \infty} \frac{1}{t_0} \int_0^{t_0} x(t, u)\, dt = 0 \tag{160}$$

if and only if

$$\lim \frac{1}{t_0} \int_0^{t_0} (u \,|\, R(t) \,|\, u)\, dt = 0, \tag{161}$$

and if and only if the spectral distribution $(u \,|\, \sigma(\lambda) \,|\, u)$ has no jump at $\lambda = 0$

$$[\sigma(+0) - \sigma(-0)]\, u = 0. \tag{162}$$

In the terminology of [Uhlhorn, 1960 b] $x(0, u)$ is an ergodic element of the Hilbert space of the stochastic process.

The conditions (160), (161), (162) are obviously much weaker than ergodicity, but may be sufficient for certain physical applications.

12. Considerations on time smoothing and smoothing of the spectrum

The values of macroscopic observables occurring in the statistical mechanics of non equilibrium phenomena are usually interpreted as averages over an interval of time necessary for the measurement. The physical reason for this averaging or time smoothing may be formulated as follows. The observable A is represented by a definite phase function or operator in the state-vector space of the mechanical system. When it is stated that the value of the observable A at the time t is A' this means that a measurement has been made in a time interval covering t, and we do not know to which time t' in this interval the value A' really belongs.

Mathematically this may be formulated in terms of a stochastic observation time with a probability distribution extending over the indicated time interval.

More precisely: the observation time is considered as the representative of a stochastic process $\tau(t)$, with stationary increments and mean value t. We shall now discuss the effect of this interpretation of the observation time on the description of series of measurements.

The representative of the observable A and the equations of motion of the mechanical system generate a Hilbert space in the usual way. On this Hilbert space the kinetics is determined by the unitary translation operator

$$U(t)=\int_{-\infty}^{+\infty}\exp(i\lambda t)\,dE_\lambda. \tag{163}$$

According to our interpretation of the observation time we have to replace $U(t)$ by the operator $U(\tau(t))$ which is random through its dependence on $\tau(t)$.

We shall be interested in the mean value of $U(\tau(t))$. We write

$$\tau(t)=t+\theta(t) \tag{164}$$

and denote the characteristic function of $\theta(t)$ by

$$\varphi(\lambda)=\mathbf{E}\exp\{i\lambda\theta(t)\} \qquad (\lambda \text{ real}). \tag{165}$$

As $\theta(t)$ is stationary $\varphi(\lambda)$ is independent of t. Hence we obtain formally the mean value of $U(\tau(t))$

$$\mathbf{E}\int_{-\infty}^{+\infty}\exp\{i(t+\theta(t))\lambda\}\,dE_\lambda=\int_{-\infty}^{+\infty}\varphi(\lambda)\exp\{it\lambda\}\,dE_\lambda=\int_{-\infty}^{+\infty}\varphi(\lambda)\,dE_\lambda\,U(t). \tag{166}$$

Hence we have obtained a linear time smoothing operator $\int_{\infty}^{+\infty}\varphi(\lambda)\,dE_\lambda$ which commutes with $U(.)$. If the probability distribution of $\theta(t)$ is derived from a density we have

$$\lim_{|\lambda|\to+\infty}\varphi(\lambda)=0 \tag{167}$$

and the effect of the time smoothing is seen to be a damping of the higher frequencies in the spectrum of $U(.)$. If the probability distribution of $\theta(t)$ is symmetric we have $\varphi(\lambda)=\varphi(-\lambda)$ and the time smoothing operator commutes with the time reversal operator T, that is time smoothing does not destroy time-reversal symmetry.

It has sometimes been suggested that time smoothing is responsible for the irreversible appearance of macroscopic phenomena.[1] Obviously this is not the case. Irreversibility and the approach to equilibrium are connected with the character of the spectral distribution of $U(t)$, in particular with the absence of jumps in this distribution and this character is obviously unaffected by time smoothing.

A smoothing of the spectrum on the other hand will always give rise to irreversibility. In its most crude form such a smoothing is achieved by the introduction of a damping factor. Sometimes a smoothing of the spectrum can be obtained as an asymptotic result when the size of the mechanical system increases infinitely. Such results can be obtained e.g. for linear mechanical systems with translation symmetry (linear chain and similar systems [Hemmer, 1959]). In this connection it must be realized that though the (point) spectrum of $U(t)$ will become denser with increasing size of the mechanical system this does not at all imply that the spectral distribution will become absolutely continuous. On the other hand a number of investigations on the approach to equilibrium of mechanical systems (van Hove, Prigogine) are based on the not easily justified assumption that the spectral distribution of $U(t)$ is absolutely continuous (or rather on assumptions equivalent to this). Hence it is not surprising that these investigations can demonstrate the existence of irreversible effects. A more thorough examination of the asymptotic spectral properties of $U(t)$ for mechanical systems of macroscopic dimensions would be of great value. However a very large number of degrees of freedoms of a mechanical system is not a necessary condition for the existence of an absolutely continuous spectral distribution. In fact one can show that the so called "Normalsysteme" which are known to be ergodic [Poincaré, 1892], [Fermi, 1923] are characterized by an absolutely continuous spectral distribution, and are hence strongly mixing in the sense of [Hopf, 1937].

---

[1] Kirkwood, lectures given at Varenna in 1959.

13. Probabilistic interpretation of the linear theory of irreversible processes

In order to obtain a probabilistic interpretation of the linear phenomenological theory of irreversible processes outlined in I ve associate with the m observables

$$A^1, \ldots, A^m \tag{168}$$

which determine the macroscopic state of a mechanical system with a stationary m-dimensional real stochastic process (section 7)

$$\{x(t, v) \,|\, -\infty < t < +\infty, \ v \in V\}. \tag{169}$$

The correspondence between the stochastic process $x(\ldots)$ and the phemenological equations is determined by the assumption that the time dependence of the conditional mean value (158) is given by the phenomenological equations I (11). With the aid of the operator $Q(t)$ introduced in I (20) this correspondence can be expressed in the form

$$R(t)^* R(0)^{-1} = Q(t) = \exp\{-tM\} \qquad (t \geqslant 0). \tag{170}$$

It is however not a priori evident that the relation (170) really defines a covariance operator. In order to show that the definition is consistent we must demonstrate that $R(t)$ defined by

$$\begin{aligned} R(t) &= R(0)\, Q(t)^* && (t > 0) \\ &= Q(-t)\, R(0) && (t \leqslant 0) \end{aligned} \tag{171}$$

satisfies the positivity condition (7).

For this purpose we determine the Fourier transform

$$\frac{1}{2\pi}\int_0^\infty R(t)\exp\{-\varepsilon t\}\exp\{-i\lambda t\}\,dt=\frac{1}{2\pi}R(0)[M^*+\varepsilon-i\lambda]^{-1} \tag{172}$$

which exists for sufficiently large $\varepsilon>0$. By the inversion theorem

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} R(0)[M^*+\varepsilon-i\lambda]^{-1}\exp\{i\lambda t\}\,d\lambda=\begin{matrix}\exp\{-\varepsilon t\}R(t) & (t>0)\\ \tfrac{1}{2}R(0) & (t=0)\\ 0 & (t<0).\end{matrix} \tag{173}$$

That is $[M^*+\varepsilon-i\lambda]^{-1}$ is a causal transform. Hence we find using (171) the following representation for $R(t)$

$$\begin{aligned}R(t)&=\lim_{\varepsilon\downarrow 0}\frac{1}{2\pi}\int_{-\infty}^{+\infty}\{R(0)[M^*-\varepsilon-i\lambda]^{-1}+[M-\varepsilon+i\lambda]^{-1}R(0)\}\exp\{i\lambda t\}\,d\lambda\\ &=\lim_{\varepsilon\downarrow 0}\frac{1}{2\pi}\int_{\infty}^{-\infty}[M+\varepsilon+i\lambda]^{-1}[M\,R(0)+R(0)\,M^*+2\,\varepsilon\,R(0)]\times\\ &\qquad\times[M^*+\varepsilon-i\lambda]^{-1}\exp\{i\lambda t\}\,d\lambda.\end{aligned} \tag{174}$$

It follows that $R(t)$ satisfies the positivity condition (7) if and only if

$$M\,R(0)+R(0)\,M^*\geqslant 0. \tag{175}$$

Comparing the condition (175) with the condition for non negative entropy production I (26) a. we find that a possible choice for $R(0)$ is

$$R(0)=\varkappa\, s^{-1}\qquad(\varkappa>0), \tag{176}$$

where s is the second order entropy operator I (20). A statistical mechanical consideration shows that, with $\varkappa$ = Boltzmann's constant,[1] this choice is correct. Hence we have found that

$$R(t)=s^{-1}Q(t)^*\qquad t\geqslant 0 \tag{177}$$

is the covariance operator of a stationary stochastic process $x(.,.)$, and its spectral distribution can be determined from (174).

The stochastic process $x(.,.)$ associated with the covariance operator $R(t)$ (177) is variant under time reversal, for the time reversal symmetry of the phenomenological equations (Onsager's reciprocal relations) is equivalent to (I, eq. (25))

$$s^{-1}Q(t)^*\,T^*=T\,Q(t)\,s^{-1} \tag{178}$$

---

[1] In the following we put Boltzmann's constant $\varkappa=1$.

and hence to
$$R(t)\,T^* = T\,R(t)^* = T\,R(-t). \tag{179}$$

This is exactly the condition (41 b). for time reversal invariance.

The relation (175) can also be obtained as a necessary condition for the non negativity of the estimation error variance (73)

$$R(0) - R(t)^*\,R(0)^{-1}\,R(t) = R(0) - Q(t)\,R(0)\,Q(t)^* \geqslant 0 \; (t \geqslant 0). \tag{180}$$

This expression is zero for $t=0$ and hence its derivative for $t=0$ must be non negative. This is exactly the condition (175). It follows from I (47) ($d = R(0)$) that the estimation error variance increases monotonically to $R(0) - d(\infty)$. In particular the phenomenological equations are strictly irreversible if and only if $d(\infty) = 0$, that is if and only if the estimation error variance reaches its maximal value in the course of time.

On the other hand the phenomenological equations are strictly irreversible if and only if the operator M has no purely imaginary eigenvalues. In this case the spectral distribution $\sigma(.)$ of $R(t)$ has a density

$$\sigma'(\lambda) = f(\lambda)^*\,f(\lambda), \qquad f(\lambda) = \frac{1}{\sqrt{2\pi}}\,[R(0)\,M^* + M\,R(0)]^{\frac{1}{2}}\,(M^* - i\lambda)^{-1}. \tag{181}$$

Obviously $J f(.)\, v$ is a causal transform (section 9). We introduce $g(.)$ in analogy to (110)

$$g(t) = \frac{1}{\sqrt{2\pi}} \int \exp\{-it\lambda\}\, f(\lambda)\, d\lambda = \begin{cases} 0 & (t > 0) \\ [R(0)\,M^* + M\,R(0)]^{\frac{1}{2}}\,Q(-t) & (t < 0), \end{cases} \tag{182}$$

and obtain a causal representation of the non deterministic process (154)

$$x(t, v) = \int_{-\infty}^{t} Z(dt', [R(0)\,M^* + M\,R(0)]^{\frac{1}{2}}\,Q(t-t')\,v), \tag{183}$$

in vector notation (section 7 eq. (35))

$$x(t) = \int_{-\infty}^{t} Q(t-t')\,[R(0)\,M^* + M\,R(0)]^{\frac{1}{2}}\,Z(dt'). \tag{184}$$

The expression (184) is the solution of the generalized Langevin equation

$$d_t\,x(t) = -M\,x(t)\,dt + [R(0)\,M^* + M\,R(0)]^{\frac{1}{2}}\,Z(dt), \tag{185}$$

that is the phenomenological equation I (11) including a purely random term. It should be observed that the Langevin equation is here obtained as a result of the assumption (171). In order to obtain the representation (183) we had to assume that M had no purely imaginary eigenvalues. As pointed out this is the necessary and sufficient condition for the phenomenological equations to determine strictly irreversible processes. Hence we find that the concepts irreversible thermodynamic process and non deterministic stochastic process correspond exactly to each other. For the existence of a purely imaginary eigenvalue of M would give rise to a jump in the spectral distribution and the stochastic process would contain a deterministic component.

We know that in general the stochastic process $x(.,.)$ is an orthogonal superposition of a deterministic and a non deterministic component. It is clear that for the particular covariance operator (177) this decomposition will correspond to the decomposition of the spectral distribution operator into an absolutely continuous and a discrete part. In order to study the general case a little closer we rewrite the equations I (46), (47) $(d = R(0))$

$$R(0) = \int_{-\infty}^{0} Q(-t')\, R(0)\, Q(-t')^* + d(\infty)$$
$$= \int_{\infty}^{0} Q(-t')\,[M\,R(0) + R(0)\,M^*]\,Q(-t')^*\,dt' + d(\infty). \tag{186}$$

With
$$g(t) = \begin{matrix} [M\,R(0) + R(0)\,M^*]^{\frac{1}{2}}\,Q(-t)^* & (t \leqslant 0) \\ 0 & (t > 0), \end{matrix} \tag{187}$$

and recalling that

$$Q(t)\,d(\infty)\,Q(t)^* = d(\infty) \text{ or equivalently } M\,d(\infty) + d(\infty)\,M^* = 0, \tag{188}$$

we obtain the following representation of $R(t)$ valid for positive and negative values of t

$$R(t) = \int_{-\infty}^{-\infty} g(t')^*\,g(t'-t)\,dt' + d(\infty)\,Q(t)^*$$
$$= \int_{-\infty}^{+\infty} g(t'+t)^*\,g(t')\,dt' + Q(-t)\,d(\infty). \tag{189}$$

This is the desired decomposition

$$R_0(t) = d(\infty)\,Q(t)^*, \quad R_1(t) = \int_{-\infty}^{+\infty} g(t')\,g(t'-t)\,dt' \tag{190}$$

for evidently $R_1(t)$ has the form (148) and the spectral distribution of $R_0(t)$ is discrete as is seen from the Fourier integral representation corresponding to (174)

$$R_0(t) = \lim_{\varepsilon \downarrow 0} \frac{1}{2\pi} \int_{-\infty}^{+\infty} [M + \varepsilon + i\lambda]^{1}\, 2\,\varepsilon\, d(\infty)\,[M^* + \varepsilon - i\lambda]^{1} \exp\{i\lambda t\}\,d\lambda. \tag{191}$$

On the other hand (190) corresponds to a decomposition into a reversible and an irreversible system (I section 3).

We denote the deterministic component of the stochastic process $x(.,.)$ by $x_0(.,.)$, the non deterministic by $x_1(.,.)$. The deterministic component is characterized by its vanishing estimation error variance and hence according to (71)

$$x_0(t, v) = x_0(0, Q(t)^* v) \quad \text{or} \quad x_0(t) = Q(t)\, x_0(0). \tag{192}$$

The non deterministic component on the other hand has the causal representation (183) (184). The composed process has the form

$$\begin{aligned} x(t) &= Q(t)\, x_0(0) + \int_{-\infty}^{t} Q(t-t')\,[R(0)\,M^* + M\,R(0)]^{\frac{1}{2}}\, Z(dt'), \\ x(t) &= Q(t)\, x(0) + \int_{-\infty}^{t} Q(t-t')\,[R(0)\,M^* + M\,R(0)]^{\frac{1}{2}}\, Z(dt'), \end{aligned} \tag{193}$$

and satisfies the Langevin equation

$$\frac{d}{dt}\, x(t) = -M\, x(t) + [R(0)\,M^* + M\,R(0)]^{\frac{1}{2}}\, \frac{Z(dt)}{dt}. \tag{194}$$

We have thus seen that the phenomenological linear equations of irreversible thermodynamics imply their probabilistic interpretation almost uniquely. It should be noted that the Langevin equation (194) is here derived as a result of the phenomenological theory whereas it is usually taken as the starting point for theories of irreversible processes based on generalized Brownian motion [Onsager and Machlup, 1953] [Hashitsume, 1952, 1956].

In the Gauszian approximation the conditional probability distribution of $x(t, .)$, given $x(0, .)$, is most easily obtained from its characteristic function and the representation (193) of $x(t, .)$. We recall that the random measure $Z(.,.)$ corresponds in the Gauszian approximation to a multidimensional Gauszian process with independent increments. The characteristic function is (v real $\in V$)

$$\begin{aligned} &\mathbf{E}\{\exp\{i\,x(t, v)\} \mid x(0, .)\} \equiv C(t, v) \\ &\quad = \exp\{i\,x(0, Q(t)^* v)\}\, \exp\Big\{-\tfrac{1}{2}\int_0^t \|\,[R(0)\,M^* + M\,R(0)]^{\frac{1}{2}}\, Q(t-t')^* v\,\|^2\, dt'\Big\} \\ &\quad = \exp\{i\,x(0, Q(t)^* v) - \tfrac{1}{2}(v \mid R(0) - Q(t)\,R(0)\,Q(t)^* \mid v)\}. \end{aligned} \tag{195}$$

It is seen that the characteristic function $C(t, v)$ satisfies the partial differential equation (in coordinate representation, comp. I.)

$$\left[\frac{\partial}{\partial t} + L^{\lambda\mu} S_{\mu\nu} V_\lambda \frac{\partial}{\partial V_\nu} + \tfrac{1}{2}(L^{\mu\nu} + L^{\nu\mu}) V_\mu V_\nu\right] C(t, v) = 0 \tag{196}$$

which is the Fourier transform of the Fokker-Planck or master-equation and can be used as a natural starting point for an extension of the statistical theory of irreversible processes to non linear phenomenological equations.

APPENDIX 1

Radon-Nikodym decomposition of a spectral measure

Definition [Halmos, 1951]: If Ω is a set with a specified σ-field $\mathfrak{F}$ of subsets, a spectral measure in Ω is a function $E(.)$ whose domain is $\mathfrak{F}$ and whose values are projections on a Hilbert space $\underline{H}$, such that $E(\Omega) = 1$ and such that $E(\bigcup_n F_n) = \sum_n E(F_n)$ whenever the F_n form a disjoint sequence of sets belonging to $\mathfrak{F}$.

Let $\mu(.)$ be a non negative σ-finite measure in Ω.

For any element x in $\underline{H}$, $\| E(.)\, x \|^2$ is a normalized measure in Ω and may according to the Radon-Nikodym theorem be uniquely decomposed into a μ-continuous and a μ-singular part

$$\| E(d\omega)\, x \|^2 = \varphi_x(\omega)\, \mu(d\omega) + \mu_x(d\omega), \tag{A1: 1}$$

where $\varphi_x(.)$ is μ-integrable and $\mu_x(.)$ vanishes outside a μ-nullset N_x.

We define the operation E^s on H by

$$E^s x = E(N_x)\, x. \tag{A1: 2}$$

The operation E^s is linear, for

$$\begin{aligned} E^s(x+y) = E(N_{x+y})(x+y) = E(N_x \cup N_y)(x+y) = E(N_x \cup N_y)\, x + \\ + E(N_x \cup N_y)\, y = E^s x + E^s y. \end{aligned} \tag{A1: 3}$$

The second step of (A1:3) follows from the triangular inequality

$$\| E(.)(x+y) \| \leqslant \| E(.)\, x \| + \| E(.)\, y \|, \tag{A1: 4}$$

showing that $\| E(.)(x+y) \|$ vanishes on any μ-nullset outside $N_x \cup N_y$. The operator E^s is now directly seen to be a projection which commutes with any $E(F)$

$$E^s E(F)\, x = E(N_x \cap F)\, x = E(F)\, E^s x. \tag{A1: 5}$$

Defining the spectral measures $E^s(.)$ and $E^a(.)$ by

$$E^s(F) = E^s E(F); \quad E^a(F) = E(F) - E^s(F), \tag{A1: 6}$$

we have succeeded in decomposing the spectral measure $E(.)$ into a μ-continuous and a μ-singular part. The two parts are orthogonal. If N is a μ-nullset we have:

$$\mu(N) = 0 \text{ implies } E^a(N) = 0. \tag{A1: 7}$$

If the Hilbert space $E^s \underline{H}$ is separable $E^s(.)$ vanishes outside a μ-nullset, for let $\{x_\nu\}$ be a sequence which is dense in $E^s \underline{H}$, and take $N = \bigcup_{\nu=1}^{\infty} N_{x_\nu}$, then $\mu(N) = 0$. Let $x \in E^s \underline{H}$ and $\| x - x_{\nu'} \| < \varepsilon$, then $\| E(N)(x - x_{\nu'}) \| = \| E(N) x - x_{\nu'} \| < \varepsilon$ and $\| x - E(N) x \| < 2\varepsilon$, that is $E(N) x = x$ for any element x in $E^s \underline{H}$ and

$$E^s = E^s(N) = E(N). \tag{A1: 8}$$

APPENDIX 2

A study of the Weyl relation

1. Let $\{U(t) \mid -\infty < t < +\infty\}$ and $\{V(t) \mid -\infty < t < +\infty\}$ be two one-parameter groups of unitary transformations on a Hilbert space $\underline{H}$, such that for arbitrary elements x and y in $\underline{H}$, $\langle y \mid U(t) \mid x \rangle$ and $\langle y \mid V(t) \mid x \rangle$ are continuous functions of t. Let the two groups be connected by the commutation relation

$$U(s)\, V(t)\, U(s)^*\, V(t)^* = \exp\{-i s t\}. \tag{A2: 1}$$

2. According to Stone's theorem the spectral decompositions of the two groups are

$$U(s) = \int_{-\infty}^{+\infty} \exp\{i s t\}\, d E_t, \tag{A2: 2}$$

$$V(s) = \int_{-\infty}^{+\infty} \exp\{i s t\}\, d F_t, \tag{A2: 3}$$

with uniquely determined spectral distributions $E.$ and $F.$ on the real axis. That is for arbitrary real t, E_t, and F_t are projection operators on $\underline{H}$ satisfying

$$E_{t_1} E_{t_2} = E_{\min(t_1, t_2)}; \ \lim_{t \to -\infty} E_t = 0; \ \lim_{t \to +\infty} E_t = 1; \ \lim_{h \downarrow 0} E_{t+h} = E_t$$

and the corresponding relations for $F.$.

3. The spectral distributions $E.$ and $F.$ satisfy the relations

$$U(s)\, F_t = F_{t+s}\, U(s), \tag{A2: 4}$$

$$E_t\, V(s) = V(s)\, E_{t+s}. \tag{A2: 5}$$

Proof: For arbitrary x and y in $\underline{H}$ we have, according to (A2: 1),

$$\langle y \mid U(s)\, V(t) \mid x \rangle = \langle y \mid V(t)\, U(s) \mid x \rangle \exp\{-i s t\}. \tag{A2: 6}$$

Introducing the spectral decomposition (A2: 2) of $U(s)$ in (A2: 6) we obtain

$$\int_{-\infty}^{+\infty} \exp\{isu\}\, d_u \langle y | E_u V(t) | x \rangle = \int_{-\infty}^{+\infty} \exp\{is(u-t)\}\, d_u \langle y | V(t) E_u | x \rangle$$
$$= \int_{-\infty}^{+\infty} \exp\{isu\}\, d_u \langle y | V(t) E_{u+t} | x \rangle. \qquad \text{(A2: 7)}$$

By application of the inversion formula for the Fourier-Stieltjes transform to (A2: 7) the equality (A2: 4) follows. The equality (A2: 5) follows from (A2: 4) and the symmetry of (A2: 1).

4. The spectral distributions $E.$ and $F.$ are absolutely continuous with respect to the Lebesgue measure on the real axis.

Proof: Let $E(.)$ and $F(.)$ be the spectral measures on the Borel field of R^1 determined by $E.$ and $F.$. That is if I is the interval $(a < t \leqslant b)$

$$E(I) = E_b - E_a; \qquad F(I) = F_b - F_a. \qquad \text{(A2: 8)}$$

It follows from (4) that if I_a^b is an arbitrary finite interval

$$F(I_a^b)\, U(t)\, F(I_a^b) = F(I_a^b)\, F(I_{a+t}^{b+t})\, U(t) = 0 \ \text{ for } |t| > (b-a). \qquad \text{(A2: 9)}$$

Hence for an arbitrary element $x \in \underline{H}$ according to (A2: 2)

$$\int_{-\infty}^{-\infty} \exp\{i\lambda t\}\, d_\lambda \langle F(I_a^b)\, x | E_\lambda | F(I_a^b)\, x \rangle = 0 \ \text{ for } |t| > (b-a). \qquad \text{(A2: 10)}$$

This is possible only if $\langle F(I_a^b)\, x | E_\lambda | F(I_a^b)\, x \rangle = \| E_\lambda F(I_a^b)\, x \|^2$ is absolutely continuous

$$\int_{-\infty}^{\lambda_0} \frac{d}{d\lambda} \| E_\lambda F(I_a^b)\, x \|^2\, d\lambda = \| E_{\lambda_0} F(I_a^b)\, x \|^2. \qquad \text{(A2: 11)}$$

Let E^s be the projection corresponding to the singular part of the spectral measure $E(.)$ (appendix 1). Equation (A2: 11) is equivalent to

$$\| E^s F(I_a^b)\, x \| = 0 \ \text{ or } \ E^s F_b\, x = E^s F_a\, x \qquad \text{(A2: 12)}$$

and this relation holds for any finite interval I_a^b. The limits of (A2: 12) for $b \to +\infty$, $a \to -\infty$ exist and hence $E^s x = 0$ for any $x \in \underline{H}$ so that we have found $E^s = 0$. From the symmetry of (A2: 1) it follows that also $F(.)$ is absolutely continuous and the proposition is proved.

5. The spectral distributions $E.$ and $F.$ have no gaps, that is

$$t \neq t' \ \text{ implies } \ E_t \neq E_{t'} \ \text{ and } \ F_t \neq F_{t'}. \qquad \text{(A2: 13)}$$

Proof: If for some $t \neq t'$ $E_t = E_{t'}$ it follows from (A2: 5) and the monotonic character of E_t as a function of t, that E_t is independent of t. Hence $E_t = 0$ and $E_t = 1$ which is a contradiction (see section 3).

6. Every non zero element $x_0 \in \underline{H}$ determines a subspace $\underline{H}[x_0]$ which is separable and invariant under both $U(.)$ and $V(.)$. The linear manifold spanned by all (finite) linear combinations of elements of the form

$$U(t)\, V(t')\, x_0 \qquad (t, t' \in R^1) \tag{A2: 14}$$

is obviously on account of (A2: 1) invariant under $U(.)$ and $V(.)$. Closing this manifold we obtain a subspace which is invariant and separable due to the fact that $U(t)$ and $V(t)$ are continuous functions of t.

Hence it is no essential restriction to assume that the Hilbert space $\underline{H}$ is separable when studying the relation (A2: 1).

7. If the Hilbert space $\underline{H}$ is separable, and if in Weyl's relation (A2: 1) $V(.)$ can be replaced by the group $V_0(.)$, then $V(.)$ and $V_0(.)$ are connected by a unitary transformation W, independent of t,

$$V_0(t) = W\, V(t)\, W^* \tag{A2: 15}$$

which commutes with all transformations of the group $U(.)$. Conversely, if W is a unitary transformation commuting with $U(.)$ the group $V_0(.)$ defined by (15) will evidently satisfy Weyl's relation.

Proof: In order to prove the first statement we shall derive an explicit representation of the Hilbert space $\underline{H}$ and of the groups $U(.)$ and $V(.)$. If g_1 is an arbitrary non zero element in $\underline{H}$ a subspace, invariant under $U(.)$ is generated by all elements which are limits of elements of the form

$$\sum_t c(t)\, U(t)\, g_1 \tag{A2: 16}$$

where $c(t')$ is an arbitrary complex function which is equal to zero for all but a finite number of values of its argument. In general this subspace is a proper subspace of $\underline{H}$. If so we can find a non zero element g_2 which is orthogonal to the subspace generated by g_1 and $U(.)$ and a second subspace generated by g_2 and $U(.)$ is obtained. This procedure may be continued in an obvious way and as the Hilbert space $\underline{H}$ was assumed to be separable we obtain finally a $U(.)$-invariant decomposition of $\underline{H}$ into finitely or denumerably many mutually orthogonal subspaces, generated by the elements

$$g_1, g_2, \ldots, g_n, \ldots \quad ; \tag{A2: 17}$$

It is possible to choose the generating elements (A2: 17) such that for any Borel set $\Delta \subset R^1$

$$E(\Delta)\, g_n = 0 \text{ if and only if } E(\Delta) = 0, \tag{A2: 18}$$

for if this condition if not originally satisfied we normalize the elements (A2: 17) so that

$$h=\sum_{n=1}^{n=\infty} g_n; \qquad \|h\|^2=\sum_{n=1}^{n=\infty}\|g_n\|^2<\infty \tag{A2: 19}$$

is convergent. It follows at once that (A2: 18) is satisfied with g_n replaced by h, for $E(\Delta)h=0$ implies $E(\Delta)g_n=0$ for every n and hence $E(\Delta)x=0$ for every x in $\underline{H}$. We may take h as the first generating element and then continue the procedure described above. In the following we assume that the g_n (A2: 17) have been chosen so as to satisfy (A2: 18). Denoting by $\underline{H}_n$ the subspace generated by g_n and $U(.)$, we have for any $x\in\underline{H}$ a unique decomposition

$$x=\sum_{n=1}^{n=\infty} x_n \qquad (x_n\in\underline{H}_n) \tag{A2: 20}$$

and for each component x_n we have (see e.g. [Achieser and Glasman, 1954]) a representation

$$x_n=\int_{-\infty}^{+\infty} x(n,\lambda)\,d\,E_\lambda\,g_n, \tag{A2: 21}$$

where $x(n,\lambda)$ is a complex function, square integrable over R^1 with respect to the measure $d_\lambda\|E_\lambda g_n\|^2$ such that with $y\in\underline{H}$

$$\langle x|y\rangle=\sum_{n=1}^{n=\infty}\langle x_n|y_n\rangle=\sum_{n=1}^{n=\infty}\int_{-\infty}^{+\infty}\overline{x(n,\lambda)}\,y(n,\lambda)\,d_\lambda\|E_\lambda g_n\|^2. \tag{A2: 22}$$

The spectral distribution $E.$ is absolutely continuous so that we can write

$$d_\lambda\|E_\lambda g_n\|^2=|f(n,\lambda)|^2\,d\lambda. \tag{A2: 23}$$

As the spectral distribution $E.$ has no gaps and as g_n satisfies (A2: 18) we have

$$f(n,\lambda)\neq 0 \text{ (for almost every } \lambda). \tag{A2: 24}$$

We define for any $x\in\underline{H}$

$$\langle n,\lambda|x\rangle\equiv x(n,\lambda)\,f(n,\lambda);\ \langle n,\lambda|x\rangle=\overline{\langle x|n,\lambda\rangle}. \tag{A2: 25}$$

It follows from (A2: 22) that

$$\langle x|y\rangle=\sum_{n=1}^{n=\infty}\int_{-\infty}^{+\infty}\langle x|n,\lambda\rangle\,d\lambda\,\langle n,\lambda|y\rangle. \tag{A2: 26}$$

For every fixed n (A2: 25) represents in view of (A2: 24) an isometric biunique correspondence between the subspace $\underline{H}_n$ of $\underline{H}$ and the Hilbert space $L^2(-\infty,+\infty)$ of functions square integrable over the real axis with respect to the Lebesgue measure $d\lambda$.

As $E.$ is the spectral distribution corresponding to $U(.)$ (A2:2) we have

$$\langle n, \lambda | U(t) | x \rangle = \exp\{i\lambda t\} \langle n, \lambda | x \rangle \text{ (for almost every } \lambda) \tag{A2:27}$$

or symbolically

$$U(t) | n, \lambda \rangle = \exp\{i\lambda t\} | n, \lambda \rangle. \tag{A2:28}$$

We note that $\langle n, \lambda |$ may be considered as a mapping of $\underline{H}_n$ onto $L^2(-\infty, \infty)$ or as a generalized eigenvector of $U(t)$ in the sense of [Dirac, 1947] with the symbolic orthogonalization condition

$$\langle n, \lambda | n', \lambda' \rangle = \delta_{n,n'} \delta(\lambda - \lambda'), \tag{A2:29}$$

which has the proper meaning (A2:26). The mapping (A2:25) is determined by the isometry condition (A2:26) up to the phase of

$$f(n, \lambda) = \langle n, \lambda | g_n \rangle. \tag{A2:30}$$

We are now going to determine the $\langle n, \lambda |$-representative of $V(t)x$ for an arbitrary $x \in \underline{H}$, $\langle n, \lambda | V(t) | x \rangle$. We start by representing $V(t) g_n$ in the form (A2:20), (A2:21)

$$V(t) g_n = \sum_{m=1}^{m=\infty} \int_{\infty}^{+\infty} v(\lambda, \lambda + t; m, n) \, d E_\lambda g_m \tag{A2:31}$$

where $v(\lambda, \lambda + t; m, n)$ is a complex function, square integrable with respect to $d_\lambda \| E_\lambda g_m \|^2$ for fixed m, and such that

$$\sum_{m=1}^{m=\infty} \int_{\infty}^{+\infty} | v(\lambda, \lambda + t; m, n) |^2 d_\lambda \| E_\lambda g_m \|^2 < +\infty.$$

The representative of $V(t)x$ is now obtained by use of the commutation relation (A2:5); with y arbitrary

$$\begin{aligned} \langle y | V(t) | x \rangle &= \sum_{m=1}^{m=\infty} \int_{\infty}^{+\infty} x(m, \lambda) \, d_\lambda \langle y | V(t) E_\lambda | g_m \rangle \\ &= \sum_{m=1}^{m=\infty} \int_{-\infty}^{+\infty} x(m, \lambda + t) \, d_\lambda \langle y | E_\lambda V(t) | g_m \rangle \\ &= \sum_{m=1}^{m=\infty} \int_{-\infty}^{+\infty} x(m, \lambda + t) \sum_{n=1}^{n=\infty} v(\lambda, \lambda + t; n, m) \, d_\lambda \langle y | E_\lambda | g_n \rangle. \end{aligned} \tag{A2:32}$$

The order of summation may be interchanged and as y is arbitrary we find

$$\langle n, \lambda \,|\, V(t) \,|\, x\rangle = \sum_{m=1}^{m=\infty} v(\lambda, \lambda+t; n, m) \langle m, \lambda+t \,|\, x\rangle. \tag{A2: 33}$$

In order to obtain the representative of $V(t)^* x$ we consider

$$\begin{aligned}\langle y \,|\, V(t)^* \,|\, x\rangle = \langle V(t)\, y \,|\, x\rangle &= \sum_{m,\,n=1}^{m,\,n=+\infty}\sum \int_{-\infty}^{+\infty} d\lambda \,\langle y \,|\, \lambda+t\rangle\, \overline{v(\lambda, \lambda+t; n, m)}\, \langle n, \lambda \,|\, x\rangle,\\ &= \sum_{m=1}^{m=\infty}\sum_{n=1}^{n=\infty} \int_{-\infty}^{+\infty} d\lambda\, \langle y \,|\, n, \lambda\rangle\, \overline{v(\lambda-t, \lambda; m, n)}\, \langle m, \lambda-t \,|\, x\rangle,\\ &= \sum_{n=1}^{n=\infty} \int_{-\infty}^{+\infty} d\lambda\, \langle y \,|\, n, \lambda\rangle \langle n, \lambda \,|\, V(t)^* \,|\, x\rangle.\end{aligned}$$

Hence
$$\langle n, \lambda \,|\, V(t)^* \,|\, x\rangle = \overline{v(\lambda-t, \lambda; m, n)}\, \langle m, \lambda-t \,|\, x\rangle. \tag{A2: 34}$$

and as $V(t)^* = V(-t)$

$$v(\lambda, \lambda-t; n, m) = \overline{v(\lambda-t; \lambda; m, n)}. \tag{A2: 35}$$

The group property of $V(.)$

$$V(t)\, V(t') = V(t+t') \tag{A2: 36}$$

gives rise to the following relation for the representative $v(\lambda, \lambda'; m, m')$

$$\sum_{m'=1}^{m'=\infty} v(\lambda, \lambda'; m, m')\, v(\lambda', \lambda''; m', m'') = v(\lambda, \lambda''; m, m'') \tag{A2: 37}$$

as is easily verified. Denoting by $\mathbf{v}(\lambda, \lambda')$ the matrix

$$\{v(\lambda, \lambda'; m, n) \,|\, m, n = 1, 2, \ldots\}$$

we can express the group property (A2: 37) and the unitarity (A2: 35) in the form

$$\begin{aligned}&\mathbf{v}(\lambda, \lambda')\,\mathbf{v}(\lambda', \lambda'') = \mathbf{v}(\lambda, \lambda''); \;\; \mathbf{v}(\lambda, \lambda) = 1,\\ &\mathbf{v}(\lambda, \lambda')^* = \mathbf{v}(\lambda', \lambda) = \mathbf{v}(\lambda, \lambda')^{-1}.\end{aligned} \tag{A2: 38}$$

With λ_0 an arbitrary fixed real number we can write

$$\mathbf{v}(\lambda, \lambda') = \mathbf{v}(\lambda_0, \lambda)^*\, \mathbf{v}(\lambda_0, \lambda'). \tag{A2: 39}$$

Further we may define a unitary transformation W on $\underline{H}$ by

$$\langle n, \lambda \,|\, W \,|\, x\rangle = \sum_{m=1}^{m=\infty} v(\lambda_0, \lambda; n, m) \langle m, \lambda \,|\, x\rangle. \tag{A2: 40}$$

It is evident that W and hence W^* commute with all transformations of the group $U(.)$. Further according to the construction

$$\langle n, \lambda \,|\, W\,V(t)\,W^* \,|\, x\rangle = \langle n, \lambda + t \,|\, x\rangle, \qquad \text{(A2: 41)}$$

showing that $W\,V(t)\,W^*$ is represented by the group of translations, when $U(t)$ is represented by the multiplicative operation (A2: 27). Otherwise expressed: by a proper choice of the generating elements (A2: 17), $U(.)$ is represented by the operation (A2: 27) and $V(.)^*$ by the group of translations. We have only to replace the generating elements $g_1, g_2, \ldots,$ by $W^* g_1, W^* g_2, \ldots$ in order to obtain the desired representation.

To make the proof of the proposition 7 complete, let $V_0(.)$ satisfy Weyl's relation (A2: 1). There will then exist a unitary W_0 such that

$$W_0\,V_0(t)\,W_0^* = W\,V(t)\,W^*$$

is represented by the group of translations, and hence

$$V_0(t) = W_0^*\,W\,V(t)\,(W_0^*\,W)^*$$

as was to be proved.

8. The groups $U(.)$ and $V(.)$ are unitarily equivalent

$$U(t) = S\,V(t)\,S^*, \qquad (-\infty < t < +\infty), \; S^{-1} = S^* \qquad \text{(A2: 42)}$$

Proof: As by a proper choice of generating elements we have

$$\langle n, \lambda \,|\, U(t) \,|\, x\rangle = \exp\{i\lambda t\}\,\langle n, \lambda \,|\, x\rangle \text{ and } \langle n, \lambda \,|\, V(t) \,|\, x\rangle = \langle n, \lambda + t \,|\, x\rangle$$

we can define S as the Fourier-Plancherel transformation (section 9)

$$\langle n, \lambda \,|\, S \,|\, x\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\{-i\lambda t\}\,dt\,\langle n, t \,|\, x\rangle. \qquad \text{(A2: 43)}$$

9. A natural generalization of Weyl's relation which has applications in physics as well as in the theory of stochastic processes is obtained by considering a representation of a locally compact, commutative group G (see e.g. [Pontrjagin, 1946]) by unitary transformations $\{U(t) \,|\, t \in G\}$ in a Hilbert space $\underline{H}$ such that $\langle x \,|\, U(t) \,|\, v\rangle$ is a continuous function of the group element t. A character $e(.)$ of the group G is by definition a complex, continuous function on G satisfying

$$e(t_1)\,e(t_2) = e(t_1 t_2) \qquad |e(t)| = 1. \qquad \text{(A2: 44)}$$

The characters of G form the commutative character group G_c, with the composition

$$e_1(t)\,e_2(t) = e_1 e_2(t). \qquad \text{(A2: 45)}$$

With G the character group G_c is locally compact. The relation between G and G_c is symmetric. Let $\{V(e) \mid e \in G_c\}$ be a representation of the character group by unitary transformations which are continuous functions of their argument, and let the groups $U(.)$ and $V(.)$ be connected by the generalized Weyl relation

$$V(e)\, U(t)\, V(e)^{-1}\, U(t)^{-1} = e(t). \tag{A2: 46}$$

It can be shown, though we shall not give the proofs, that most of the propositions about the Weyl relation (A2: 1) discussed above remain true in a generalized sense for the relation (A2: 46). In particular there are essentially two representations of the relation (A2: 46), one in terms of the Hilbert space $L^2(G)$ of complex functions on G, square integrable with respect to the invariant (Haar) measure $\mu(.)$, and one in terms of the Hilbert space $L^2(G_c)$ corresponding to the character group. Hence we will obtain a decomposition of the Hilbert space $\underline{H}$ into orthogonal subspaces each invariant under $U(.)$ and $V(.)$ and isometrically equivalent to $L^2(G)$ and to $L^2(G_c)$. In fact if x is an element of one such subspace and if its $L^2(G)$-representative is $\langle t;\, G \mid x\rangle$, its $L^2(G_c)$-representative $\langle e;\, G_c \mid x\rangle$ we have the generalized Fourier transformation

$$\langle e;\, G_c \mid x\rangle = \int_G e(t)\, \langle t;\, G \mid x\rangle\, \mu(dt), \tag{A2: 47}$$

$$\langle t;\, G \mid x\rangle = \int_{G_c} \overline{e(t)}\, \langle e;\, G_c \mid x\rangle\, \mu_c(de), \tag{A2: 48}$$

$$\|x\|^2 = \int_G |\langle t;\, G \mid x\rangle|^2\, \mu(dt) = \int_{G_c} |\langle e;\, G_c \mid x\rangle|^2\, \mu_c(dt). \tag{A2: 49}$$

The equation (A2: 49) is the general form of Parseval's relation. Further the representations of $U(.)$ and $V(.)$ are given by

$$\langle t;\, G \mid U(t_1) \mid x\rangle = \langle t_1^{-1} t;\, G \mid x\rangle, \tag{A2: 50}$$

$$\langle t;\, G \mid V(e) \mid x\rangle = e(t)\, \langle t;\, G \mid x\rangle, \tag{A2: 51}$$

and hence by application of (A2: 47), (A2: 48)

$$\langle e;\, G_c \mid U(t) \mid x\rangle = e(t_1)\, \langle e;\, G_c \mid x\rangle, \tag{A2: 52}$$

$$\langle e;\, G_c \mid V(e_1) \mid x\rangle = \langle e_1 e;\, G_c \mid x\rangle. \tag{A2 53}$$

Possible and important examples of G are: the group of translations in a finite-dimensional Euclidean space, corresponding to a finite set of canonical coordinates in quantum mechanics, the group of integers with addition as composition, the group of translations mod 2π etc.

ACKNOWLEDGEMENT

I wish to thank professor Oskar Klein for his kind interest in the present work and constant encouragement.

Institute for theoretical physics, University of Sockholm, Stockholm. February 1960.

REFERENCES

Achieser, A. I., and Glasmann, I. M., 1954, Theorie der linearen Operatoren im Hilbertschen Raum, Akademie-Verlag.

Dirac, P. A. M., 1947, The principles of quantum mechanics, Oxford University Press.

Fermi, E., 1923, Beweis, dass ein mechanisches Normalsystem im allgemeinen quasi-ergodisch ist, Phys. Zeitschr. *24*, 261.

Gibbs, J. W., 1902, Elementary principles in statistical mechanics, Charles Scribner's Sons.

Grenander, U., 1950, Stochastic processes and statistical inference, Arkiv för Matematik, *1*, 195.

Grenander, U., and Szegö, G., 1958, Toeplitz forms and their applications, University of California Press.

Halmos, P., 1951, Introduction to Hilbert space and the theory of spectral multiplicity, Chelsea.

Hanner, O., 1950, Deterministic and non-deterministic stationary random processes, Arkiv för Matematik, *1*, 161.

Hashitsume, N., 1952, 1956, A statistical theory of linear dissipative systems, Progr. Th. Phys. *8*, 461, *15*, 369.

Hemmer, P. Chr., 1959, Dynamic and stochastic types of motion in the linear chain, Thesis, Trondheim.

Hopf, E., 1937, Ergodentheorie, Springer.

van Hove, L., 1957, The approach to equilibrium in quantum statistics, Physica *22*, 441.

Jaglom, A. M., 1959, Einführung in die Theorie der stationären Zufallsfunktionen, Akademie-Verlag.

Karhunen, K., 1947, Über lineare Methoden in der Wahrscheinlichkeitsrechnung, Annales Academiae Scientiarum Fennicae A I 37.

—— 1950, Über die Struktur stationärer zufälliger Funktionen, Arkiv för Matematik, *1*, 141.

von Neumann, J., 1931, Die Eindeutigkeit der Schrödingerschen Operatoren, Math. Annalen *104*, 570.

Onsager, L., and Machlup, S., 1953, Fluctuations and irreversible processes, Phys. Rev. *91*, 1505, 1512.

Poincaré, H., 1892. Les méthodes nouvelles de la mécanique céleste, tome I, Gauthiers-Villars.

Pontrjagin, L., 1946, Topological groups, Princeton University Press.

Titchmarsh, E. C., 1948, Introduction to the theory of Fourier integrals, Oxford University Press.

Toll, J. S., 1956, Causality and the dispersion relation: logical foundations, Phys. Rev. *104*, 1760.

Uhlhorn, U., 1960 I, On the foundations of the linear theory of irreversible processes I, Arkiv för Fysik *17*, 257.

—— 1960 a, Macroscopic observables and generalized canonical ensembles, Arkiv för Fysik *17*, 233.

—— 1960 b, On statistical mechanics of non equilibrium phenomena, Arkiv för Fysik *17*, 193.

Weyl, H., 1927, Quantenmechanik und Gruppentheorie, Zeitschr. Physik *46*, 1.

—— 1931, Gruppentheorie und Quantenmechanik. Hirzel.

Wold, H., 1938, A study in the analysis of stationary time series, Almqvist & Wiksell.

Tryckt den 4 april 1960

Uppsala 1960. Almqvist & Wiksells Boktryckeri AB